



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 124533**

**TO: Bao-Qun Li**  
**Location: REM-3D24&3C18**  
**Art Unit: 1648**  
**Thursday, June 17, 2004**

**Case Serial Number: 09/899303**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**Remsen 1A59**  
**Phone: 571-272-2523**

**toby.port@uspto.gov**

### **Search Notes**

Dear Examiner Li,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1958.63 Seconds  
(without alignments)  
14206.949 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCGGTGTCTTTCTC.....TACTCTTGTCTCTTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pin.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	642	100.0	642	6	A48665	A48665 Sequence 3
2	642	100.0	642	6	AR157324	AR157324 Sequence
3	642	100.0	642	6	AX452752	AX452752 Sequence
4	642	100.0	642	6	AX685004	AX685004 Sequence
5	642	100.0	642	6	AX794846	AX794846 Sequence
6	628.2	97.9	795	6	A48667	A48667 Sequence 5
7	628.2	97.9	795	6	AR157325	AR157325 Sequence
8	628.2	97.9	795	6	AX452754	AX452754 Sequence
9	628.2	97.9	795	6	AX685006	AX685006 Sequence
10	628.2	97.9	795	6	AX794848	AX794848 Sequence
11	619.8	96.5	2082	6	A48709	A48709 Sequence 47
12	619.8	96.5	2082	6	AR157350	AR157350 Sequence
13	619.8	96.5	2082	6	AX452796	AX452796 Sequence
14	619.8	96.5	2082	6	AX685048	AX685048 Sequence
15	619.8	96.5	2082	6	AX794890	AX794890 Sequence
16	619.8	96.5	2433	6	A48711	A48711 Sequence 49
17	619.8	96.5	2433	6	AR157351	AR157351 Sequence
18	619.8	96.5	2433	6	AX452798	AX452798 Sequence
19	619.8	96.5	2433	6	AX685050	AX685050 Sequence
20	619.8	96.5	2433	6	AX794892	AX794892 Sequence
21	575	89.6	9379	14	AF207766	AF207766 Hepatitis
22	573.4	89.3	9410	14	HPCKLR2	D50481 Hepatitis C
23	569.6	88.7	9418	14	HCVL32996	AU132996 Hepatitis
24	568.6	88.6	9386	14	AF165056	AF165056 Hepatitis
25	567	88.3	9386	14	AF165055	AF165055 Hepatitis
26	567	88.3	9410	14	HPCKLS2	D50485 Hepatitis C
27	567	88.3	9598	14	AB049101	AB049101 Hepatitis
28	566.4	88.2	3461	6	E07544	E07544 Gene fragme
29	566.4	88.2	3461	6	E09288	E09288 cDNA encodi
30	565.4	88.1	1880	14	HPCSTRJ4	D00832 Hepatitis C
31	565.4	88.1	2540	6	E04260	E04260 cDNA encodi
32	565.4	88.1	2540	6	E04805	E04805 cDNA to 5'
33	565.4	88.1	2540	6	E07391	E07391 cDNA encodi
34	565.4	88.1	3296	14	AB008441	AB008441 Hepatitis
35	565.4	88.1	9033	14	HCUJ238800	AU238800 Hepatitis
36	565.4	88.1	9448	14	HPCJ483	D13558 Hepatitis C
37	565.4	88.1	9456	14	HPCRNA	D10934 Hepatitis C
38	565.4	88.1	9605	6	AX739971	AX739971 Sequence
39	565.4	88.1	9605	14	HCUJ238799	AU238799 Hepatitis
40	565.4	88.1	11076	6	AR406040	AR406040 Sequence
41	565.4	88.1	11076	6	AR406046	AR406046 Sequence
42	565.4	88.1	11076	6	AR406048	AR406048 Sequence
43	565.4	88.1	11076	6	AR406050	AR406050 Sequence
44	565.4	88.1	11076	6	AX036252	AX036252 Sequence
45	565.4	88.1	11076	6	AX036258	AX036258 Sequence

ALIGNMENTS

RESULT 1

A48665

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

A48665 Sequence 3 from Patent WO9604385.  
A48665  
A48665.1 GI:2302378

unidentified

unclassified

1 (bases 1 to 642)

Maertens,G., Bosman,F., De,M.G. and Buysse,M.

PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND

THERAPEUTIC USE

Patent: WO 9604385-A 3 15-FEB-1996;

linear PAT 07-MAR-1997



```
VERSION      AX452752.1  GI:21712437
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Hepatitis C virus
              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens, G., Bosman, P., de Martynoff, G. and Buyse, M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 3 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..642
                  /organism="Hepatitis C virus"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:11103"
                1..639
                  /note="unnamed protein product"
                  /codon_start=1
                  /protein_id="CAD38066.1"
                  /db_xref="GI:21712438"
                  /db_xref="REMTREMBL:CAD38066"
                  /translation="MPGCSFSIFELLALLSCLTIPASAYEVNRVNSGMVHTNDCNSSI
VIEADMIHTPGCVPCVRENNSRCWVALTPTLAARNASVPTTIRHVDLLVGA
LCSAMTVGLCGSVFLVQLFTSPRRHETVQDCNCSYPGHITGHRMADMMNNSP
TTALVVSQLLRIPOAVVDMVAGAHGVLGLAYYSVNGNWKVLLVLLFAL"
              mat_peptide
                1..636
                  /product="unnamed"
              CDS
                100.0%; Score 642; DB 6; Length 642;
                Best Local Similarity 100.0%; Pred. No. 1.4e-150;
                Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              1  ATGCCCGGTGCTCTTCTCTATCTTCTCTGCTTACTGCTTACTGCTCTGCTGACCAATCCA 60
              2  ATGCCCGGTGCTCTTCTCTATCTTCTCTGCTTACTGCTTACTGCTCTGCTGACCAATCCA 60
              3  GCTTCCGCTTATGAGTGCGGACGCTGTCGGGATGACCATGTACGAAAGCTGCTCC 120
              4  GCTTCCGCTTATGAGTGCGGACGCTGTCGGGATGACCATGTACGAAAGCTGCTCC 120
              5  AACTCAAGCATTTGTATGAGGACGACATCATGATCATGACACACCCCGGCTGCTGCTCC 180
              6  AACTCAAGCATTTGTATGAGGACGACATCATGATCATGACACACCCCGGCTGCTGCTCC 180
              7  TGCGTTCCGGAGAACAACTCTTCCCGCTGCTGCGGATGCTGCTGCTGCTGCTGCTGCT 240
              8  TGCGTTCCGGAGAACAACTCTTCCCGCTGCTGCGGATGCTGCTGCTGCTGCTGCTGCT 240
              9  AGGAACGCCAGCTGCCACACGACATGACGACGACGACGACGACGACGACGACGACGAC 300
              10 AGGAACGCCAGCTGCCACACGACATGACGACGACGACGACGACGACGACGACGACGAC 300
              11 GCTGCTCTCTGTTCCGCTATGATGAGTGGGAGTCTTCCCGATCTGCTGCTGCTGCTGCT 360
              12 GCTGCTCTCTGTTCCGCTATGATGAGTGGGAGTCTTCCCGATCTGCTGCTGCTGCTGCT 360
              13 CAGCTGTTTCAACATCTCGCTCGCCGATGAGACGCTGAGACGCTGAGACGCTGAGAC 420
              14 CAGCTGTTTCAACATCTCGCTCGCCGATGAGACGCTGAGACGCTGAGACGCTGAGAC 420
              15 TATCCCGGACATATACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 480
              16 TATCCCGGACATATACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 480
              17 ACAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
              18 ACAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
              19 GCGGGGGCCATGTTGGGGAGTCTTGGGGGCTGCGGCTTACTTATTCATGCTGCTGCTG 600
              20 GCGGGGGCCATGTTGGGGAGTCTTGGGGGCTGCGGCTTACTTATTCATGCTGCTGCTG 600

RESULT 4
AX685004
LOCUS       AX685004             642 bp      DNA      linear      PAT 29-MAR-2003
DEFINITION Sequence 3 from Patent WO02055548.
ACCESSION   AX685004
VERSION     AX685004.1  GI:29371409
KEYWORDS    .
SOURCE      Hepatitis C virus
ORGANISM    Hepatitis C virus
              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens, G., Bosman, F. and Buyse, M.A.
TITLE        Purified Hepatitis C Virus envelope proteins for diagnostic and
              therapeutic use
JOURNAL      Patent: WO 02055548-A 3 18-JUL-2002;
              INNOGENETICS N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..642
                  /organism="Hepatitis C virus"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:11103"
                1..639
                  /note="unnamed protein product"
                  /codon_start=1
                  /protein_id="CAD86519.1"
                  /db_xref="GI:29371410"
                  /db_xref="REMTREMBL:CAD86519"
                  /translation="MPGCSFSIFELLALLSCLTIPASAYEVNRVNSGMVHTNDCNSSI
VIEADMIHTPGCVPCVRENNSRCWVALTPTLAARNASVPTTIRHVDLLVGA
LCSAMTVGLCGSVFLVQLFTSPRRHETVQDCNCSYPGHITGHRMADMMNNSP
TTALVVSQLLRIPOAVVDMVAGAHGVLGLAYYSVNGNWKVLLVLLFAL"
              mat_peptide
                1..636
                  /product="unnamed"
              ORIGIN
                Query Match      100.0%; Score 642; DB 6; Length 642;
                Best Local Similarity 100.0%; Pred. No. 1.4e-150;
                Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              1  ATGCCCGGTGCTCTTCTCTATCTTCTCTGCTTACTGCTTACTGCTCTGCTGACCAATCCA 60
              2  ATGCCCGGTGCTCTTCTCTATCTTCTCTGCTTACTGCTTACTGCTCTGCTGACCAATCCA 60
              3  GCTTCCGCTTATGAGTGCGGACGCTGTCGGGATGACCATGTACGAAAGCTGCTCC 120
              4  GCTTCCGCTTATGAGTGCGGACGCTGTCGGGATGACCATGTACGAAAGCTGCTCC 120
              5  AACTCAAGCATTTGTATGAGGACGACATCATGATCATGACACACCCCGGCTGCTGCTCC 180
              6  AACTCAAGCATTTGTATGAGGACGACATCATGATCATGACACACCCCGGCTGCTGCTCC 180
              7  TGCGTTCCGGAGAACAACTCTTCCCGCTGCTGCGGATGCTGCTGCTGCTGCTGCTGCT 240
              8  TGCGTTCCGGAGAACAACTCTTCCCGCTGCTGCGGATGCTGCTGCTGCTGCTGCTGCT 240
              9  AGGAACGCCAGCTGCCACACGACATGACGACGACGACGACGACGACGACGACGACGAC 300
              10 AGGAACGCCAGCTGCCACACGACATGACGACGACGACGACGACGACGACGACGACGAC 300
              11 GCTGCTCTCTGTTCCGCTATGATGAGTGGGAGTCTTCCCGATCTGCTGCTGCTGCTGCT 360
              12 GCTGCTCTCTGTTCCGCTATGATGAGTGGGAGTCTTCCCGATCTGCTGCTGCTGCTGCT 360
              13 CAGCTGTTTCAACATCTCGCTCGCCGATGAGACGCTGAGACGCTGAGACGCTGAGAC 420
              14 CAGCTGTTTCAACATCTCGCTCGCCGATGAGACGCTGAGACGCTGAGACGCTGAGAC 420
              15 TATCCCGGACATATACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 480
              16 TATCCCGGACATATACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 480
              17 ACAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
              18 ACAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
              19 GCGGGGGCCATGTTGGGGAGTCTTGGGGGCTGCGGCTTACTTATTCATGCTGCTGCTG 600
              20 GCGGGGGCCATGTTGGGGAGTCTTGGGGGCTGCGGCTTACTTATTCATGCTGCTGCTG 600
```



```
QY 421 TATCCCGCCACATAAACAGGTACCGTATGGCTTCGGGATATGATGATGAACGTGGTGGCT 480
DB 421 TATCCCGCCACATAAACAGGTACCGTATGGCTTCGGGATATGATGATGAACGTGGTGGCT 480
QY 481 ACAACGGCCCTGGTGGTATGCGAGCTGTCGGGATCCCAACAGCTGTCGTGGACATGGTG 540
DB 481 ACAACGGCCCTGGTGGTATGCGAGCTGTCGGGATCCCAACAGCTGTCGTGGACATGGTG 540
QY 541 GCGGGGGCCCATTCGGGAGTCTCGGCGGCTCGGCTACTATTCATGCTGGGGAACCTGG 600
DB 541 GCGGGGGCCCATTCGGGAGTCTCGGCGGCTCGGCTACTATTCATGCTGGGGAACCTGG 600
QY 601 GCTAAGGTTTTCATGATGATGATCTCTTTGGCTCTCTAATAG 642
DB 601 GCTAAGGTTTTCATGATGATGATCTCTTTGGCTCTCTAATAG 642

RESULT 5
AX794846
LOCUS AX794846 642 bp DNA linear PAT 04-OCT-2003
DEFINITION Sequence 3 from Patent WO03051912.
ACCESSION AX794846
VERSION AX794846.1 GI:37515735
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
SOURCE Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Depla,E. and Bosman,P.
TITLE Purified Hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL INNOCENTICS N.V. (BE)
PATENT: WO 03051912-A 3 26-JUN-2003;
INNOGENETICS N.V. (BE)
FEATURES
Location/Qualifiers
1..642
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:11103"
1..639
/organism="unassigned protein product"
/codon_start=1
/protein_id="CAE48242.1"
/db_xref="GI:37515736"
/translation="*WPGSPSIFILLALLSCLTIPASAYEVNVSVMHVTNDCNSSI
VYEADIMHTPGVPCVRENNSRCWALPTLAARNASVPTTIRHVDLVGAHA
LCSAMYVDLCSGPLVLSQPTISPRHETVQDCNCSYPGHITGHRMADMMNWS
TTALVVSQLLRIPQAVDMVAGAHGVLGLAYTSMVGNWAKVLIVMLLPAL"
mat_peptide 1..636
/product="unnamed"
CDS
Query Match 100.0%; Score 642; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.4e-150;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCCGGTTGCTCTTCTCTATCTTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCT 60
DB 1 ATGCCCGGTTGCTCTTCTCTATCTTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCT 60
QY 61 GCTTCGCTTATGAGTGGCGCAACGTGTCGGGATGTACCATGTCCAGAACGACTGCTCC 120
DB 61 GCTTCGCTTATGAGTGGCGCAACGTGTCGGGATGTACCATGTCCAGAACGACTGCTCC 120
QY 121 AACTCAAGCATTTGTATGAGGACGACGACATGATCATGCACACCCCGGGTGGTGGCC 180
DB 121 AACTCAAGCATTTGTATGAGGACGACGACATGATCATGCACACCCCGGGTGGTGGCC 180
QY 181 TCGGTTCCGGGAGAACATCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGAGCT 240
DB 181 TCGGTTCCGGGAGAACATCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGAGCT 240
```

```
QY 241 AGGAACGCGCAGCTCCCCACACCAACAATACGACGCCAGCTCGATTGCTCGTTGGGGCG 300
DB 241 AGGAACGCGCAGCTCCCCACACCAACAATACGACGCCAGCTCGATTGCTCGTTGGGGCG 300
QY 301 GCTGCTCTCTGTTTCGGCTATGTACGTGGGGATCTCTGCGGATCTGCTCTCTCTGCTCTCC 360
DB 301 GCTGCTCTCTGTTTCGGCTATGTACGTGGGGATCTCTGCGGATCTGCTCTCTCTGCTCTCC 360
QY 361 CAGCTGTTTCAACCATCTCGCTCGCGCATGAGCGGTGACGAGCTGCAATTTGCTCAATC 420
DB 361 CAGCTGTTTCAACCATCTCGCTCGCGCATGAGCGGTGACGAGCTGCAATTTGCTCAATC 420
QY 421 TATCCCGCCACATAAACAGGTACCGTATGGCTTCGGGATATGATGATGAACGTGGTGGCT 480
DB 421 TATCCCGCCACATAAACAGGTACCGTATGGCTTCGGGATATGATGATGAACGTGGTGGCT 480
QY 481 ACAACGGCCCTGGTGGTATGCGAGCTGTCGGGATCCCAACAGCTGTCGTGGACATGGTG 540
DB 481 ACAACGGCCCTGGTGGTATGCGAGCTGTCGGGATCCCAACAGCTGTCGTGGACATGGTG 540
QY 541 GCGGGGGCCCATTCGGGAGTCTCGGCGGCTCGGCTACTATTCATGCTGGGGAACCTGG 600
DB 541 GCGGGGGCCCATTCGGGAGTCTCGGCGGCTCGGCTACTATTCATGCTGGGGAACCTGG 600
QY 601 GCTAAGGTTTTCATGATGATGATCTCTTTGGCTCTCTAATAG 642
DB 601 GCTAAGGTTTTCATGATGATGATCTCTTTGGCTCTCTAATAG 642

RESULT 6
AX48667
LOCUS AX48667 795 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 5 from Patent WO9604385.
ACCESSION AX48667
VERSION AX48667.1 GI:2302380
KEYWORDS unidentified
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 795)
AUTHORS Maertens,G., Bosman,P., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL INNOCENTICS NV (BE)
PATENT: WO-9604385-A 5-15-FEB-1996;
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
Location/Qualifiers
1..795
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..792
/organism="unassigned protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302381"
/db_xref="REMTREMBL:CAA03129"
/translation="*MLGRVDTLTCGPADLVGYIPLVGPALGGAALAHGVRLVLEDG
VNYATGNLPGCSFISIFLLALLSCLTVPASAYEVNVSVMHVTNDCNSSI VTEAADM
IMHTPGVPCVRENNSRCWALPTLAARNASVPTTIRHVDLVGLVGAHAFCSAMV
GDLCSGPLVLSQPTISPRHETVQDCNCSYPGHITGHRMADMMNWSPTTALVVS
QLLRIPQAVDMVAGAHGVLGLAYTSMVGNWAKVLIVMLLPAP"
mat_peptide 1..789
/product="unnamed"
ORIGIN
Query Match 97.9%; Score 628.2; DB 6; Length 795;
Best Local Similarity 98.8%; Pred. No. 4e-147;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TCCCGGTTGCTCTTCTCTATCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
```





ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.

1  
REFERENCE  
AUTHORS Maertens, G., Depla, E. and Bosman, F.  
TITLE Purified Hepatitis C virus envelope proteins for diagnostic and  
therapeutic use  
JOURNAL Patent: WO 03051912-A 5 26-JUN-2003;  
INNOGENETICS N.V. (BE)  
FEATURES  
source Location/Qualifiers  
1..795  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
1..792  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB48243.1"  
/db\_xref="GI:37515738"  
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEDG  
VNYATGNLPGCSFSIFLLALLSCLTPASAYEVNVSVMYHVTNDSCNSIVYBAADM  
IMHTPGVPCVRENSSRCWVALLPTLAARNASVPTTIRRHVDLLVGAARFCSAMTV  
GDLGSGFLVLSQFTISPRHETVQDCNCSYIPGHITGRMAMDMNMMNSPTTALVVS  
QLLRIPQAVDMVAGAHGVLAYYSYMGVNAKVLIVMLLPAP"  
mat\_peptide 1..789  
/product="unnamed"

ORIGIN  
Query Match 97.9%; Score 628.2; DB 6; Length 795;  
Best Local Similarity 98.8%; Pred. No. 4e-147;  
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTTCTTGGCTTACTGTCTGTGACCAATTCAG 61  
DB 155 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTGGTCTGTCTGTGACCGTTCAG 214  
QY 62 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 121  
DB 215 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 274  
QY 122 ACTCAGCATTGTGTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 181  
DB 275 ACTCAGCATTGTGTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 334  
QY 182 ACTCAGCATTGTGTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 241  
DB 335 GCGTTCGGGAGACAACTCTTCCCGCTGCTCGGATGCGCTCACCCCGGTCGCTCCA 394  
QY 242 GCGTTCGGGAGACAACTCTTCCCGCTGCTCGGATGCGCTCACCCCGGTCGCTCCA 301  
DB 602 CTAAGGTTTGAATGTGATGCTACTCTTTGCTCTCTAATAG 642

Db 755 CTAAGGTTTGAATGTGATGCTACTCTTTGCTCCTAATAG 795

RESULT 11  
LOCUS A48709 2082 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 47 from Patent WO9604385.  
ACCESSION A48709  
VERSION A48709.1 GI:2302422  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 2082)  
AUTHORS Maertens, G., Bosman, F., De, M. G. and Buysse, M.  
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND  
THERAPEUTIC USE  
JOURNAL Patent: WO 9604385-A 47 15-FEB-1996;  
INNOGENETICS NV (BE)  
COMMENT Other publication CA 2172273 960215  
Other publication AU 3382495 960304.  
FEATURES  
source Location/Qualifiers  
1..2082  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
1..2079  
/note="unnamed protein product; Protein sequence is in  
conflict with the conceptual translation"  
/codon\_start=1  
/protein\_id="CAA03146.1"  
/db\_xref="GI:2302423"  
/db\_xref="REMBL:CAA03146"  
/translation="NLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEDG  
VNYATGNLPGCSFSIFLLALLSCLTPASAYEVNVSVMYHVTNDSCNSIVYBAADM  
IMHTPGVPCVRENSSRCWVALLPTLAARNASVPTTIRRHVDLLVGAARFCSAMTV  
GDLGSGFLVLSQFTISPRHETVQDCNCSYIPGHITGRMAMDMNMMNSPTTALVVS  
QLLRIPQAVDMVAGAHGVLAYYSYMGVNAKVLIVMLLPAP"  
mat\_peptide 1..2076  
/product="unnamed"

ORIGIN  
Query Match 96.5%; Score 619.8; DB 6; Length 2082;  
Best Local Similarity 98.9%; Pred. No. 5.3e-145;  
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTCTTCTTGGCTTTACTGTCTGTGACCAATTCAG 61  
DB 155 TGCCCGGTGCTCTTCTCTATCTCTTCTTGGCTTTGCTGTCTGTGACCGTTCAG 214  
QY 62 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 121  
DB 215 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 274  
QY 122 ACTCAGCATTGTGTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 181  
DB 275 ACTCAGCATTGTGTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACATGCTCCA 334  
QY 182 GCGTTCGGGAGACAACTCTTCCCGCTGCTCGGATGCGCTCACCCCGGTCGCTCCA 241  
DB 335 GCGTTCGGGAGACAACTCTTCCCGCTGCTCGGATGCGCTCACCCCGGTCGCTCCA 394  
QY 242 GCGTTCGGGAGACAACTCTTCCCGCTGCTCGGATGCGCTCACCCCGGTCGCTCCA 301



[illegible]

RESULT 14					
AX685048					
LOCUS	AX685048	2082 bp	DNA	linear	PAT 29-MAR-2003
DEFINITION	Sequence 47 from Patent WO0205548.				
ACCESSION	AX685048				
VERSION	AX685048.1	GI:29371453			
KEYWORDS	Hepatitis C virus				
SOURCE	Hepatitis C virus				
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				

REFERENCE	1
AUTHORS	Maertens, G., Bosman, F. and Buyse, M.A.
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL	Patent: WO 02055548-A 47 18-JUL-2002;

```

FEATURES
source
    Location/Qualifiers
      1..2082
        /organism="Hepatitis C virus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:11103"
      <1..2079
        /notes="unnamed protein product"
      /codon_start=1
      /protein_id="CAD86535.1"
      /db_xref="GI:29371454"
      /db_xref="RENTREMBL:CAD86535"
      /translations="NLGKVIQDTITCGCFADLVGVIPLVGLGPGCAARALAHGVRLIEDG

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 241.553 Seconds  
(without alignments)

11290.892 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGGCTCTTCTC.....TACTCTTTGGCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002as.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	642	100.0	642	2	AAT12704 HCV E1 CO
2	642	100.0	642	6	AAL48913 Hepatitis
3	642	100.0	642	9	Add55634 Hepatitis
4	628.2	97.9	795	2	AAT12705 HCV E1 CO
5	628.2	97.9	795	6	AAL48914 Hepatitis
6	628.2	97.9	795	9	Add55513 Hepatitis
7	619.8	96.5	2082	6	AAL48939 Hepatitis
8	619.8	96.5	2082	9	Add55555 Hepatitis
9	619.8	96.5	2086	2	AAT12973 HCV E1 CO
10	619.8	96.5	2433	2	AAT12974 HCV E1 CO
11	619.8	96.5	2433	9	Add55557 Hepatitis
12	608.8	94.8	2434	6	AAL48940 Hepatitis
13	566.4	88.2	3461	2	AAG64068 Non-A, no
14	566.4	88.2	3461	2	AAT30386 5'UTR/COR
15	565.4	88.1	1880	2	AAG24467 NANB hepa
16	565.4	88.1	2187	2	ABA03491 Cuticle p
17	565.4	88.1	2540	2	AAG43889 NANB hepa
18	565.4	88.1	2540	2	AAQ63753 NANBHV ge
19	565.4	88.1	9605	6	ABK91431 Hepatitis
20	565.4	88.1	9605	6	ABK91424 Hepatitis
21	565.4	88.1	9605	6	ABK91429 Hepatitis
22	565.4	88.1	9605	6	ABK91432 Hepatitis
23	565.4	88.1	9605	6	ABK91411 Hepatitis

## ALIGNMENTS

RESULT 1  
AAT12704

ID AAT12704 standard; DNA; 642 BP.

AC AAT12704;

DT 23-SEP-1996 (first entry)

XX HCV E1 construct HCC19A.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX 88.

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, De Martynoff G, Buysse M;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins

XX - in presence of disulphide bond cleavage agent, to produce proteins

XX suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1

XX and E2 protein coding sequence constructs. These sequences are included

XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.

XX The recombinant proteins can then be isolated using a method of the

XX invention. In the method, the envelope proteins are purified by carrying

XX out a disulphide bond cleavage, or a reduction step with a disulphide

XX bond cleavage agent, after lysis of recombinant host cells. The

XX constructs containing the purified HCV envelope proteins can be used for

XX vaccinating humans against HCV, for in vitro detection of HCV antibodies

XX in a sample, and in a serotyping assay for detecting one or more

24	565.4	88.1	9605	6	ABK91430	Hepatitis
25	565.4	88.1	9605	6	ABK91428	Hepatitis
26	565.4	88.1	9605	6	ABK91425	Hepatitis
27	565.4	88.1	9605	6	ABK91426	Hepatitis
28	565.4	88.1	9605	6	ABK91433	Hepatitis
29	565.4	88.1	9605	6	ABK91427	Hepatitis
30	565.4	88.1	9608	6	ABK91427	Hepatitis
31	565.4	88.1	11062	6	AAQ25331	Hepatitis
32	565.4	88.1	11076	3	AAQ98965	Hepatitis
33	563.8	87.8	1863	2	AAQ15363	Pragment
34	563.2	87.7	1251	2	AAQ26981	HCV gene
35	563.2	87.7	3360	2	AAT03677	Hepatitis
36	563.2	87.7	9413	2	AAQ80498	DNA encod
37	563.2	87.7	9413	2	AAQ81559	Hepatitis
38	563.2	87.7	9413	2	AAQ03960	Partial H
39	563.2	87.7	9413	6	AAQ25517	Hepatitis
40	563.2	87.7	9413	7	AAL53723	Hepatitis
41	563.2	87.7	9413	7	AAQ49655	Hepatitis
42	562.2	87.6	2540	2	AAQ29628	Hepatitis
43	560	87.2	1251	2	AAQ25610	HCV in ex
44	560	87.2	1398	2	AAQ22140	Hepatitis
45	559	87.1	742	2	AAQ20926	C10-E15 D





```

Db      421 TATCCCGCCACATAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGGTGGCT 480
Qy      481 ACAACGGCCCTGGTGTATGCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Db      481 ACAACGGCCCTGGTGTATGCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Qy      541 GCGGGGGCCCATTTGGGGAGTCTGCGGGCCCTCGCTACTATTCCATGGTGGGAACTGG 600
Db      541 GCGGGGGCCCATTTGGGGAGTCTGCGGGCCCTCGCTACTATTCCATGGTGGGAACTGG 600
Qy      601 GCTAAGGTTTGGATTTGATGCTACTCTTTTGCTCTCTAATAG 642
Db      601 GCTAAGGTTTGGATTTGATGCTACTCTTTTGCTCTCTAATAG 642

RESULT 3
ADD55634
ID      ADD55634 standard; DNA; 642 BP.
XX
AC      ADD55634;
XX
DT      15-JAN-2004 (first entry)
XX
Hepatitis C virus E1 protein coding sequence #1.
XX
Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
KW      liver fibrosis; ds; gene.
XX
Hepatitis C virus.
OS
PN      W02003051912-A2.
XX
PD      26-JUN-2003.
XX
PF      18-DEC-2002; 2002WO-EP014480.
XX
PR      18-DEC-2001; 2001US-00020510.
XX
PR      16-OCT-2002; 2002US-0418358P.
XX
PA      (INNO-) INNOGENETICS NV.
XX
PI      Maertens G, Depla E, Bosman F;
XX
WPI; 2003-541632/51.
DR
P-PSDB; ADD55512.
XX
New hepatitis C virus (HCV) vaccine composition, useful for reducing
PT      liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX
Example 1; SEQ ID NO 3; 271pp; English.
XX
The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
CC      liver disease. The vaccine of the invention comprises an HCV E1 or E2
CC      protein as an antigen. The HCV vaccine is useful for reducing liver
CC      disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
CC      present DNA sequence encodes an HCV E1 protein.
XX
SQ      Sequence 642 BP; 109 A; 195 C; 176 G; 162 T; 0 U; 0 Other;

Query Match      100.0%; Score 642; DB 9; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.2e-174;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGCCCGGTGCTCTTTCTATCTCTCTGCTTTGCTTTACTGCTGCTGACCATTTCCA 60
Db      1 ATGCCCGGTGCTCTTTCTATCTCTCTGCTTTGCTTTACTGCTGCTGACCATTTCCA 60
Qy      61 GCTTCGCTTATGAGTGGCGCAAGCTGTCCGGGATGTACCATGTACGAAAGCACTGTCC 120
Db      61 GCTTCGCTTATGAGTGGCGCAAGCTGTCCGGGATGTACCATGTACGAAAGCACTGTCC 120
Qy      121 AACTCAAGCATTTGTGTATGAGCGGAGCATGATCATGCACACCCCGGGTGGTGGCC 180

```

```

Db      121 AACTCAAGCATTTGTGTATGAGCGGAGCATGATCATGCACACCCCGGGTGGTGGCC 180
Qy      181 TCGGTTTCGGAGAACAACTCTTTCCGCTGCTGGTAGCGCTCACCCGCCAGCTCGCAGCT 240
Db      181 TCGGTTTCGGAGAACAACTCTTTCCGCTGCTGGTAGCGCTCACCCGCCAGCTCGCAGCT 240
Qy      241 AGGAACGCGAGGTCCTCCACACGACAAATACGACCGCAGCTGATTTGCTTGGGGGG 300
Db      241 AGGAACGCGAGGTCCTCCACACGACAAATACGACCGCAGCTGATTTGCTTGGGGGG 300
Qy      301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTCGGATCTGCTCTTCCCTCGTCTCC 360
Db      301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTCGGATCTGCTCTTCCCTCGTCTCC 360
Qy      361 CAGCTGTTTACCATCTCGCTCGCGGCATGAGAAGCGTGCAGGACTGCAATTTGCTCAATC 420
Db      361 CAGCTGTTTACCATCTCGCTCGCGGCATGAGAAGCGTGCAGGACTGCAATTTGCTCAATC 420
Qy      421 TATCCCGGCCACATAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGGTGGCT 480
Db      421 TATCCCGGCCACATAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGGTGGCT 480
Qy      481 ACAACGGCCCTGGTGTATGCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Db      481 ACAACGGCCCTGGTGTATGCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Qy      541 GCGGGGGCCCATTTGGGGAGTCTGCGGGCCCTCGCTACTATTCCATGGTGGGAACTGG 600
Db      541 GCGGGGGCCCATTTGGGGAGTCTGCGGGCCCTCGCTACTATTCCATGGTGGGAACTGG 600
Qy      601 GCTAAGGTTTGGATTTGATGCTACTCTTTTGCTCTCTAATAG 642
Db      601 GCTAAGGTTTGGATTTGATGCTACTCTTTTGCTCTCTAATAG 642

RESULT 4
AAT12705
ID      AAT12705 standard; DNA; 795 BP.
XX
AC      AAT12705;
XX
DT      23-SEP-1996 (first entry)
XX
HCV E1 construct HCC110A.
DE
KW      HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW      serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW      ss.
XX
Hepatitis C virus.
OS
PN      W09604385-A2.
XX
PD      15-FEB-1996.
XX
PP      31-JUL-1995; 95WO-EP003031.
XX
PR      29-JUL-1994; 94EP-00870132.
XX
(INNO-) INNOGENETICS NV.
Maertens G, Bosman F, De Martynoff G, Buyse M;
WPI, 1996-129401/13.
Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
- in presence of d,l-sulphide bond cleavage agent, to produce proteins
suitable for direct use in vaccines or diagnostic assays of HCV.
Claim 23; Fig 21; 146pp; English.
AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included

```













QY 242 GGAACGCGAGCGTCCCAACCAACGACAAATACGACGCGCAAGTTCGATTTGCTCGTTGGGGCGG 301  
DB GGAACGCGAGCGTCCCAACCAACGACAAATACGACGCGCAAGTTCGATTTGCTCGTTGGGGCGG 805  
QY 302 CTGCTCTCTGTTCCGCTATGTAAGTGGGGGATCTCTGGGATCTCTCTCTCTCTCTCTCTCTCC 361  
DB CTGCTCTCTGTTCCGCTATGTAAGTGGGGGATCTCTGGGATCTCTCTCTCTCTCTCTCTCTCC 865  
QY 362 AGCTGTTTCCACATCTCGCTCGCGGATGAGACGCTGACGAGTGCATTTGCTCAATCT 421  
DB AGCTGTTTCCACATCTCGCTCGCGGATGAGACGCTGACGAGTGCATTTGCTCAATCT 925  
QY 422 ATCCCGGCGACATTAACAGGTCAACGATGCTTGGGATATGATGAATGATGCTGCTGCTGCTA 481  
DB ATCCCGGCGACATTAACAGGTCAACGATGCTTGGGATATGATGAATGATGCTGCTGCTGCTA 985  
QY 482 CAACGGCCCTGCTGATCGCAGCTGCTCCGATCCCAACAGCTGCTGCGACATGCTGG 541  
DB CAACGGCCCTGCTGATCGCAGCTGCTCCGATCCCAACAGCTGCTGCGACATGCTGG 1045  
QY 542 CGGGGGCCCATTTGGGAGTCTCTGGGGGCTCTGCTACTATTCCATGCTGCTGCTGCTGCTG 601  
DB CGGGGGCCCATTTGGGAGTCTCTGGGGGCTCTGCTACTATTCCATGCTGCTGCTGCTGCTG 1046  
QY 602 CTAAAGTTTGTGATGCTACTCTTTTCG 632  
DB CTAAAGTTTGTGATGCTACTCTTTTCG 1136

## RESULT 12

AA048940  
ID AAL48940 standard; DNA; 2434 BP.  
AC AAL48940;  
XX  
DT 24-OCT-2002 (first entry)  
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.  
DE  
KM Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;  
KM virucide; immunostimulant; vaccine; ds.  
XX  
OS Hepatitis C virus.  
XX  
PN W0200255548-A2.  
XX  
PD 18-JUL-2002.  
XX  
PF 11-JAN-2002; 2002MO-EP000219.  
XX  
PR 11-JAN-2001; 2001US-0260669P.  
XX  
PR 30-AUG-2001; 2001US-0315768P.  
XX  
PA (INNO-) INNOGENETICS NV.  
XX  
PI Maertens G, Boeman F, Buysse M;  
XX  
DR WPI; 2002-599657/64.  
DR P-PSDB; AAO18679.

PT New therapeutic vaccine compositions comprising at least one purified  
PT recombinant hepatitis C virus (HCV) single or specific oligomeric  
PT recombinant envelope protein E1 or E2, useful for immunizing humans from  
PT HCV infection.

PS Example 2; Page 212-215; 243pp; English.

XX The present invention relates to new therapeutic vaccine compositions for  
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a  
XX composition containing at least one purified recombinant HCV single or  
XX specific oligomeric recombinant envelope proteins selected from an E1 and  
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are

CC useful for inducing HCV-specific antibodies or for immunizing humans  
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as  
CC vaccines or therapeutics, in HCV screening and confirmatory antibody  
CC tests, for raising antibodies, in the preparation of medicament, and for  
CC in vitro monitoring of HCV disease or prognosing the response to  
CC treatment of patients suffering from HCV infection. The present sequence  
CC is a coding sequence described in the exemplification of the invention  
XX

Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 U; 0 Other;

Query Match 94.8%; Score 608.8; DB 6; Length 2434;  
Best Local Similarity 98.7%; Pred. No. 1.3e-164;  
Matches 624; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 2 TCCCGGTTGCTCTTCA 60  
DB TCCCGGTTGCTCTTCA 565  
QY 61 GCTTCGCTTATGAGGTGCGCAACGTTGCCGGATGATACCATGTCCAGAACGACTGCTCC 120  
DB GCTTCGCTTATGAGGTGCGCAACGTTGCCGGATGATACCATGTCCAGAACGACTGCTCC 625  
QY 121 AACTCAACGATTTGTATGAGCGACGACATGATGACACACCGGGTGGTGGCC 180  
DB AACTCAACGATTTGTATGAGCGACGACATGATGACACACCGGGTGGTGGCC 685  
QY 181 TGGTTTGGGAGAACAACTCTTCCCGTCTGCTGGGTAGCGCTCACCCCGCTCCAGCT 240  
DB TGGTTTGGGAGAACAACTCTTCCCGTCTGCTGGGTAGCGCTCACCCCGCTCCAGCT 745  
QY 241 AGGAACGCGACGCTCCCAACCAACGACAAATACGACGCGCAATGCTGATTTGCTTGGGGCG 300  
DB AGGAACGCGACGCTCCCAACCAACGACAAATACGACGCGCAATGCTGATTTGCTTGGGGCG 805  
QY 301 GCTGCTCC 360  
DB GCTGCTCC 865  
QY 361 CAGCTGTTTCAACATCTCGCTCGCGCATGACGAGCTGCAATTTGCTCAATC 420  
DB CAGCTGTTTCAACATCTCGCTCGCGCATGACGAGCTGCAATTTGCTCAATC 925  
QY 421 TATCCCGGCGCACATAACGAGTCAACGCTATGCTTTGGGATATGATGAATGCTGCTCGCT 480  
DB TATCCCGGCGCACATAACGAGTCAACGCTATGCTTTGGGATATGATGAATGCTGCTCGCT 985  
QY 481 ACAACGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB ACAACGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045  
QY 541 GCGGGGGCCCATTTGGGGGATGCTGCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 600  
DB GCGGGGGCCCATTTGGGGGATGCTGCGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 1046  
QY 601 GCTAAGGTTTGTGATGCTACTCTTTTCG 632  
DB GCTAAGGTTTGTGATGCTACTCTTTTCG 1137

## RESULT 13

AA064068  
ID AAQ64068 standard; cDNA; 3461 BP.

XX AAQ64068;  
AC AAQ64068;  
XX

DT 14-FEB-1995 (first entry)

DE Non-A, non-B hepatitis virus gene #4.

XX Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV;  
KM NS1; NS2; NS3; antigen; detection; ss.  
XX Hepatitis C virus.  
OS



```

XX Key Location/Qualifiers
FH CDS 307..3461
FT /*tag= a
FT misc_RNA 307..879
FT /*tag= b
FT /*label= core
FT misc_RNA 880..1455
FT /*tag= c
FT /*label= ENV
FT misc_RNA 1456..2736
FT /*tag= d
FT /*label= NS1
FT misc_RNA 2737..3461
FT /*tag= e
FT /*label= NS2-3
FT /*note= "NS3 N-terminal"
XX JP06141870-A.
XX 24-MAY-1994.
XX 12-MAR-1992; 92JP-00088140.
XX 12-MAR-1992; 92JP-00088140.
XX (TOKR-) ZH TOKYO RINSHO IGAKU SOGO KENKYUSHO.
XX (SANW) SANWA KAGAKU KENKYUSHO CO.
XX (TOFU) TONEN CORP.
XX WPI; 1994-205028/25.
XX P-PSDB; AAR54066.
XX DNA coding a Non-A Non-B hepatitis virus antigen - useful for detecting
XX HCV within serum.
XX Claim 1-5; Page 11-15; 22pp; Japanese.
XX Hepatitis C virus #4 and #6 genes were isolated (AA064068-69). Both genes
XX contain the core, ENV, NS1, NS2 and NS3 regions. A core region fragment
XX is given in AA064067
XX Sequence 3461 BP; 638 A; 1046 C; 1012 G; 765 T; 0 U; 0 Other;
XX Query Match 88.2%; Score 566.4; DB 2; Length 3461;
XX Best Local Similarity 93.5%; Pred. No. 2.3e-152;
XX Matches 591; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
XX 2 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTACTGCTTCTGCTGACCATCCAG 61
XX 812 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTGCTGCTGACCATCCAG 871
XX 62 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGTACCATGTACGACGACTGCTCCA 121
XX 872 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGTACCATGTACGACGACTGCTCCA 931
XX 122 ACTCAAGCATTTGTATGAGGCGAGCGACATGATCATGCACACCCCGGGTGGTGGCCCT 181
XX 932 ACTCAAGTATTGTATGAGGCGAGCGACGTATCATGCACACCCCGGGTGGTGGCCCT 991
XX 182 GCGTTCGGGAGAACAACTCTTCCGCTGCTGCGGTAGCGCTACCCCGACGCTCGAGCTA 241
XX 992 GCGTTCGGGAGAACAACTTCTCCGCTGCTGCGGTAGCGCTACCTCCAGCTCGCGCCA 1051
XX 242 GGAACGCCAGCGTCCCGACGACAAATACGACGCGCACTGCGATTTGCTGCTGGCGCG 301
XX 1052 GAAACGACGAGTCCCGACGACAAATACGACGCGCACTGCGATTTGCTGCTGGCGCG 1111
XX 302 CTGCTCTCTGTTCCGCTATGATGCGGGGATCTCTGCGGATCTGTCTTCGCTGCC 361
XX 1112 CTGCTCTCTGCTCCGCAATGATGCGGGGATCTCTGCGGATCTGTCTTCCTCGTCCC 1171
XX 362 AGCTGTTACCATCTCTGCGCTCGCCGCGCATGAGACGGTGCAGGACTGCAATTGCTCAATCT 421

```

```

Db 1172 AGCTGTTACCTTCTCACCCTGCGGTATGAGCGGTAGAGACTCAGGACTCAATCT 1231
QY 422 ATCCCGGCGCACATAACAGGTCAACCGTATGCGCTTGGGATATGATGATGAACCTGGTCCCTA 481
Db 1232 ATCCCGGCGCACAGTGTGTCAGTCAACCGTATGCGCTTGGGATATGATGATGAACCTGGTCCCTA 1291
QY 482 CAACGCGCTGCTGCTGATGCGGAGTCTGCGGATCCCAAGCTGCTGCGGACATGGTGG 541
Db 1292 CAACGCGCTGCTGCTGATGCGGAGTCTGCGGATCCCAAGCTGCTGCGGACATGGTGG 1351
QY 542 CGGCGGCGCCATTTGGGAGTCTGCGGCGGCTCTGCTTACTTCCATGCTGCTGGGAACTGGG 601
Db 1352 CAGGCGGCGCACTGCGGAGTCTGCGGCGGCTCTGCTTACTTCCATGCTGCTGGGAACTGGG 1411
QY 602 CTAAGGCTTTGATGCGGAGTCTGCTTCTTGGT 633
Db 1412 CTAAGGCTTTGATGCGGAGTCTGCTTCTTGGT 1443

RESULT 14
AAT30386
ID AAT30386 standard; cDNA; 3461 BP.
XX AC AAT30386;
XX 27-AUG-2003 (revised)
DT 22-AUG-1996 (first entry)
XX 5'UTR/CORE/ENV/NS1/NS2/NS3 cDNA from HCV (#4).
XX Hepatitis C virus; HCV; antigen; detection; antibody; ds.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FH CDS 307..3461
FT /*tag= a
FT misc_feature 307..879
FT /*tag= b
FT /*product= "Core peptide"
FT misc_feature 880..1455
FT /*tag= c
FT /*product= "ENV1"
FT misc_feature 1456..2736
FT /*tag= d
FT /*product= "NS1/ENV2"
FT misc_feature 2737..3461
FT /*tag= e
FT /*product= "NS2 and NS3"
XX JP07133291-A.
XX 23-MAY-1995.
XX 18-JUN-1993; 93JP-00147944.
XX 18-JUN-1993; 93JP-00147944.
XX (TOFU) TONEN CORP.
XX WPI; 1995-220780/29.
XX P-PSDB; AAR98361.
XX Recombinant polypeptide comprising partial NS1 region of hepatitis non-A
XX non-B viral antigen - used in a method for detecting antibodies against
XX hepatitis non-A non-B virus.
XX Disclosure; Page 10-12; 15pp; Japanese.
XX The sequences given in AAT30386-87 encode the 5'UTR/CORE/ENV/NS1/NS2/ NS3
XX protein region derived from hepatitis C virus (HCV) isolates #4 and #6
XX respectively. The proteins encoded by these sequences partic. encode
XX CC

```



Db 1327 CAACAGCCCTAGTGTGTGTCGAGTTGCTCCGATGCCACAAAGCTGTCGTGGACATGGTGG 1386  
Qy 542 CGGGGGCCCATTTGGGGAGTCTCTGGGGCCCTCGCCTACTATTCCATGGTGGGNACTGGG 601  
Db 1387 CGGGGGCCCACTGGGGAGTCTCTGGGGCCCTTGCCTACTATTCCATGGTAGGGAACTGGG 1446  
Qy 602 CTAAGGTTTTGATTGTGANGCTACTCTTTGC 632  
Db 1447 CTAAGGTCCTGATTGTGGCGTACTCTTGC 1477

Search completed: June 16, 2004, 09:36:51  
Job time : 244.553 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 1725.91 Seconds  
(without alignments)  
11108.065 Million cell updates/sec

Title: US-09-899-303A-3  
Perfect score: 642  
Sequence: 1 ATGCCGGTGGCTTTCTC.....TACTCTTGGCTCTAATAG 642

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estom:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gsal:\*

29: gb\_gss2:\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	117.2	18.3	488	9	AV755731
C 2	92.6	14.4	492	9	AV758366
C 3	41.6	6.5	534	14	CF846043
C 4	40.6	6.3	925	29	CNS0091P

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	40.4	6.3	590	14	CA659369
6	40.4	6.3	645	29	CNS01213
7	40	6.2	525	10	BB337089
8	40	6.2	671	12	BI723733
9	40	6.2	740	12	BJ536071
10	40	6.2	1201	13	BX356664
11	39.6	6.2	423	9	AA459034
12	39	6.1	621	14	CA816001
13	39	6.1	624	12	BI723734
14	38.6	6.0	399	9	AV638521
15	38.6	6.0	434	9	AV637507
16	38.6	6.0	440	9	AV637983
17	38.6	6.0	450	9	AV637259
18	38.6	6.0	451	9	AV637328
19	38.6	6.0	451	9	AV637643
20	38.6	6.0	453	9	AV634724
21	38.6	6.0	454	9	AV637050
22	38.6	6.0	456	9	AV635382
23	38.6	6.0	473	9	AV632765
24	38.6	6.0	481	9	AV635503
25	38.6	6.0	485	9	AV632811
26	38.6	6.0	506	9	AV392445
27	38.6	6.0	508	9	AV634095
28	38.6	6.0	526	9	AV641895
29	38.6	6.0	533	9	AV638125
30	38.6	6.0	537	9	AV632335
31	38.6	6.0	588	9	AV387329
32	38.6	6.0	693	13	CA160850
33	38.2	6.0	431	9	AV639153
34	38.2	6.0	501	9	AV638474
35	38.2	6.0	1148	14	CK208676
36	38	5.9	742	13	BQ752673
37	38	5.9	1195	28	E10902
38	37.8	5.9	435	13	C72860
39	37.8	5.9	634	10	BF396591
40	37.4	5.8	497	9	AV633658
41	37.4	5.8	655	29	CG274746
42	37.4	5.8	782	29	CC605275
43	37.4	5.8	856	28	BZ578381
44	37.4	5.8	860	29	CG286130
45	37.4	5.8	872	28	BZ555011

ALIGNMENTS

RESULT 1

AV755731/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AV755731 488 bp mRNA linear EST 19-OCT-2000  
AV755731 BM Homo sapiens cDNA clone BMAKB03 5', mRNA sequence.

AV755731.1 GI:10913579

EST.

Homo sapiens (human)

Homo sapiens

Sukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 488)

Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,

Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,

Zeng, B., Xu, S., Gu, W., Tu, Y., Jia, J., Pu, G., Ren, S., Zhong, M.,

Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z., and

Chen, Z.

Homo sapiens cDNA BM clones

Unpublished (2000)

Contact: Zequang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex. 45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.







## ORIGIN

Query Match 6.2%; Score 40; DB 10; Length 525;  
Best Local Similarity 47.0%; Pred. No. 50;  
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCCAACTCAAGCATGTGTATGAGCGAGCGGACATGATCATGCACACCC 167  
DB 43 GCACCGCTTCAACAACACCCCGTGTCTGAATTCGCGCGCCCTACAAAGTACCTGCTGTCCAT 102

QY 168 CGGGTGGTGGTCCCTGCGTTCGGGAGAACACTCTTCCGCTGCTGGGTAGCGTCAACCC 227  
DB 103 GGAGGACAGACCTCCACCTGCGAGGTATCGAGTACTGCGACCTGGGCAACCTGTCCAA 162

QY 228 CACGCTCGAGTAGGAAACGACGCTCCCAACACGACATACGACCCACGCTCGATTT 287  
DB 163 CGCGCTCAAAAACAACATCTTCATGATCCCAACCCGTCATCGCGGGCGGGCGC 222

QY 288 GCTGCTGGGGCGGCTGCTCTCTGTTCCGCTATGATGAGGGGATCTCTCGGATCTGT 347  
DB 223 GGGCGAGCGCGCGCGCGGAGAGCTAGCGAGCGCGCGGCGAGCGCCATGAAGT 282

QY 348 CTTCCTCGTCTCCAGCTGTTCAC 371  
DB 283 CAACATGCGCACCTGCTGTCTAC 306

## RESULT 8

BI723733 671 bp mRNA linear EST 19-SEP-2001  
LOCUS 1031067F08.y1 C. reinhardtii CC-1690, Stress II (normalized),  
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI723733

VERSION BI723733.1 GI:15699428

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii

ORGANISM Chlamydomonas reinhardtii

REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadales; Chlamydomonas.

1 (bases 1 to 671)

Grossman, A., Chang, C.-W., Davies, J., Harris, B., Hauser, C.,

Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants. Project: 1031

Unpublished (2001)

CONTACT: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1. .671

/organism="Chlamydomonas reinhardtii"

/mol\_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db\_xref="taxon:3055"

/clone\_lib="C. reinhardtii CC-1690, Stress II

(normalized), Lambda Zap II"

/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:

XhoI; Stress condition II library, constructed by John

Davies and Jeffrey McDermott, combines cDNAs from CC-1690

cells grown to mid-log phase in TAP (NH4+ - containing)

and shifted to TAP - NO3- (24hrs); H2 production

conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant

Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +

sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).

PolyA mRNA was purified from each sample, pooled and cDNA

synthesized. The cDNA was directionally cloned into lambda

zap II (Stratagene) in the EcoRI (5') and XhoI (3')

sites. pBluescript II SK- plasmids were excised from the

## ORIGIN

Query Match 6.2%; Score 40; DB 12; Length 671;  
Best Local Similarity 47.0%; Pred. No. 55;  
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCCAACTCAAGCATGTGTATGAGCGAGCGGACATGATCATGCACACCC 167  
DB 353 GCACCGCTTCAACAACACCCCGTGTCTGAATTCGCGCGCCCTACAAAGTACCTGCTGTCCAT 412

QY 168 CGGGTGGTGGTCCCTGCGTTCGGGAGAACACTCTTCCGCTGCTGGGTAGCGTCAACCC 227  
DB 413 GGAGGACAGACCTCCACCTGCGAGGTATCGAGTACTGCGACCTGGGCAACCTGTCCAA 472

QY 228 CACGCTCGAGTAGGAAACGACGCTCCCAACACGACATACGACCCACGCTCGATTT 287  
DB 473 CGCGCTCAAAAACAACATCTTTCATGATCCCAACCCGTCATCGCGGGCGGGCGC 532

QY 288 GCTGCTGGGGCGGCTGCTCTCTGTTCCGCTATGATGAGGGGATCTCTCGGATCTGT 347  
DB 533 GGGCGAGCGCGCGCGGCGGAGAGCTAGCGAGCGCGCGGCGAGCGCCATGAAGT 592

QY 348 CTTCCTCGTCTCCAGCTGTTCAC 371  
DB 593 CAACATGCGCACCTGCTGTCTAC 616

## RESULT 9

BJ536071/c

LOCUS BJ536071

DEFINITION BJ536071 MP01SSB cDNA Oryzias latipes cDNA clone MF01SSB002D03 3',

mRNA sequence.

ACCESSION BJ536071

VERSION BJ536071.1 GI:22194883

KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)

ORGANISM Oryzias latipes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;

Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

1 (bases 1 to 740)

Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.

Medaka EST Project in Takeda's lab

Unpublished (2001)

Contact: Tadasi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. .740

/organism="Oryzias latipes"

/mol\_type="mRNA"

/strain="Hd-rh"

/db\_xref="taxon:8090"

/clone="MF01SSB002D03"

/sex="mixture of female and male"

/tissue\_type="whole embryo"

/dev\_stage="segmentation stage 20 - 25"

/clone\_lib="MF01SSB cDNA"

ORIGIN

Query Match 6.2%; Score 40; DB 12; Length 740;  
Best Local Similarity 52.1%; Pred. No. 58;  
Matches 88; Conservative 0; Mismatches 81; Indels 0; Gaps 0;



QY 446 GTATGCTTGGATATGATGAACTGGTCGCTCAACAGCCCGCTGGTATCCAGC 505  
 DB 606 GTACGGCTTGGCTGGAGACTCCCTCAACCCACCTGCTGAATATGATGTCGCCCAAG 547  
 QY 506 TGCTCCGATCCACAAAGCTGCTGGACATGCTGGGGGGGCCCATTCGGAGAGCTCTCG 565  
 DB 546 TCTTTGAGACCTCGCCACATGATCCAGCGGTGATGGGGCCCTGGAGGGTCTCCGG 487  
 QY 566 CGGGCTCGCCCTACTATTCATGCTGGGGAACCTGGGCTAAGGTTTGTAT 614  
 DB 486 TGGCCCTCGCCCTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438

RESULT 10  
 BX356664  
 LOCUS  
 DEFINITION  
 clone CS0D1015YB03 3-PRIME, mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1015CA02NP1.

FEATURES  
 source  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1015YB03"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 6.2%; Score 40; DB 13; Length 1201;  
 Best Local Similarity 13.6%; Pred. No. 71;  
 Matches 77; Conservative 223; Mismatches 264; Indels 4; Gaps 1;

QY 9 TTGCTCTTCTATCTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 68  
 DB 525 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 584  
 QY 69 TTATGAGTGGCAAGCTGTCGGGATGATGATGATGATGATGATGATGATGATGAT 128  
 DB 585 TTCTTTTNN 644  
 QY 129 CATTTGATGATGAGCAGCAGCATGATCATGACACCCCGGCTGCTGCTGCTGCTGCT 188  
 DB 645 TSNVTTTBTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 704  
 QY 189 GGAGAACAACTCTCCGCTGCTGGGTAGCGCTACACCCACCCACCCACCCACCCAC 248  
 DB 705 STTSTTSBSSSTTTTTSBSSSTTSBSSSTTSBSSSTTSBSSSTTSBSSSTTSB 764  
 QY 249 CAGCGTCCCAACACACATACAGCCACGCTGATGCTGCTGCTGCTGCTGCTGCTCT 308

DB 765 STTBYTBSSSYBSBS - - - - - BTSSCSTSSBTSSTSSBTSSTSSBTSSTSSBTSST 820  
 QY 309 CTGTTCCGCTATGATGCTGGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCT 368  
 DB 821 SSTTTSTKSSBTTTSSBTSBSSTSTTTTBTCTSTTTTSSSYSSBSSTTSTSTSTTB 880  
 QY 369 CACCATCTCCCTCGCCGCGGATGAGAGGCTGAGGACTGCAATGCTCAATCTATCCGG 428  
 DB 881 SWSBBSSTTTTSTSTSTTTTSSSSSTTTTSSSSSTTTTSSSSSTTTTSSSSSTTTSSSS 940  
 QY 429 CCACATAACAGGTACCGTATGCTGCTGATGATGATGATGATGATGATGATGATGATG 488  
 DB 941 STSBTSBS 1000  
 QY 489 CCTGGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548  
 DB 1001 STTTTSSSTBS 1060  
 QY 549 CCATGGGAGTCTCGCGGCGCTCGCC 576  
 DB 1061 BBTTTTBTSTBTSSBSBSBSBSBSBSBSBSBSBSBSBSBSBSBSBSBSBSBS 1088

RESULT 11  
 AA459034  
 LOCUS  
 DEFINITION  
 aa26b11.r1 NCI\_CGAP\_GCB1 423 bp mRNA linear EST 13-AUG-1997  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 423)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb-x@mail.nih.gov  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 382.

FEATURES  
 source  
 1..423  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:6032522"  
 /db\_xref="taxon:9606"  
 /clone="IMAGR:814365"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP GCB1"  
 /notes="Vector: PT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD20+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 [5'-TGTACCATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-3',  
 1. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
 Query Match 6.2%; Score 39.6; DB 9; Length 423;  
 Best Local Similarity 47.9%; Pred. No. 57;





Search completed: June 16, 2004, 13:27:08  
Job time : 1729.91 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 46.2775 Seconds  
(without alignments)  
7698.741 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCGGTGTCTTCTC.....TACTCTTGTCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/ACTUS COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	3	US-08-612-973-3
2	642	100.0	642	3	US-08-927-597-3
3	628.2	97.9	795	3	US-08-612-973-5
4	628.2	97.9	795	3	US-08-927-597-5
5	619.8	96.5	2082	3	US-08-612-973-47
6	619.8	96.5	2082	3	US-08-927-597-47
7	619.8	96.5	2433	3	US-08-612-973-49
8	619.8	96.5	2433	3	US-08-927-597-49
9	565.4	88.1	11076	4	US-09-539-601-19
10	565.4	88.1	11076	4	US-09-539-601-19
11	565.4	88.1	11076	4	US-09-539-601-25
12	565.4	88.1	11076	4	US-09-539-601-31
13	563.8	87.8	1539	2	US-08-470-426B-17
14	563.8	87.8	1863	2	US-08-470-426B-14
15	559	87.1	742	1	US-08-081-072-18
16	559	87.1	742	1	US-08-449-093A-18
17	559	87.1	932	1	US-08-081-072-15
18	559	87.1	932	1	US-08-449-093A-15
19	555.8	86.6	2116	3	US-08-191-160-21
20	555.8	86.6	9595	3	US-09-014-416-4
21	555.8	86.6	9599	3	US-09-014-416-6
22	544.6	84.8	9472	4	US-08-150-204E-96
23	540.8	84.2	1167	1	US-08-324-977-9
24	540.8	84.2	1167	2	US-08-384-616-9
25	540.8	84.2	1167	2	US-08-904-686A-9
26	540.8	84.2	1167	3	US-09-315-850-9
27	540.8	84.2	1499	1	US-08-324-977-3

28	540.8	84.2	1499	2	US-08-384-616-3	Sequence 3, Appli
29	540.8	84.2	1499	2	US-08-904-686A-3	Sequence 3, Appli
30	540.8	84.2	1499	3	US-09-315-850-3	Sequence 3, Appli
31	540.8	84.2	6039	1	US-08-324-977-11	Sequence 11, Appl
32	540.8	84.2	6039	2	US-08-384-616-11	Sequence 11, Appl
33	540.8	84.2	6039	2	US-08-904-686A-11	Sequence 11, Appl
34	540.8	84.2	6039	3	US-09-315-850-11	Sequence 11, Appl
35	540.8	84.2	9030	1	US-08-324-977-13	Sequence 13, Appl
36	540.8	84.2	9030	2	US-08-384-616-13	Sequence 13, Appl
37	540.8	84.2	9030	2	US-08-904-686A-13	Sequence 13, Appl
38	540.8	84.2	9030	3	US-09-315-850-13	Sequence 13, Appl
39	540.8	84.2	9416	1	US-08-324-977-1	Sequence 1, Appli
40	540.8	84.2	9416	2	US-08-384-616-1	Sequence 1, Appli
41	540.8	84.2	9416	2	US-08-904-686A-1	Sequence 1, Appli
42	540.8	84.2	9416	3	US-09-315-850-1	Sequence 1, Appli
43	540.8	84.2	9416	4	US-08-823-895A-27	Sequence 27, Appl
44	511.8	79.7	576	1	US-08-086-428B-21	Sequence 21, Appl
45	511.8	79.7	576	2	US-08-468-570-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1  
US-08-612-973-3  
; Sequence 3, Application US/08612973  
; Patent No. 6150334  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, PONS  
; APPLICANT: DE MARTINOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612.973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 642 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..639  
; FEATURE:  
; NAME/KEY: mat peptide  
; LOCATION: 1..636



QY 541 GCGGGGCGCCATTGGGGAGTCTCGCGGCTCGCTACTATTCCATGCTGGGGAACCTGG 600  
Db 541 GCGGGGCGCCATTGGGGAGTCTCGCGGCTCGCTACTATTCCATGCTGGGGAACCTGG 600  
QY 601 GCTAAGGTTTGTGATGTGATGCTACTCTTTGCTCTCTTAATAG 642  
Db 601 GCTAAGGTTTGTGATGTGATGCTACTCTTTGCTCTCTTAATAG 642

## RESULT 3

US-08-612-973-5  
; Sequence 5, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 795 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..792  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..789  
; US-08-612-973-5

Query Match 97.9%; Score 628.2; DB 3; Length 795;  
Best Local Similarity 98.8%; Pred. No. 2.2e-157;  
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTTACTGTCTGCTGACCAATTCAG 61  
Db 155 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTTACTGTCTGCTGACCGTTCAG 214  
QY 62 CTTCCGCTTATGAGGTGCGCAAGCTGTCCGGGATGTACCATGTCCAGAACGACTGCTCCA 121  
Db 215 CTTCCGCTTATGAGGTGCGCAAGCTGTCCGGGATGTACCATGTCCAGAACGACTGCTCCA 274

QY 122 ACTCAAGCATTTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGCTGCGTCCCT 181  
Db 275 ACTCAAGCATTTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGCTGCGTCCCT 334  
QY 182 GCGTTTCGGGAGAACAACTCTTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTA 241  
Db 335 GCGTTTCGGGAGAACAACTCTTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTA 394  
QY 242 GGAAGCGGCGTCCCGACACGACATACAGCGCCACGATGATGCTGTTGGGCGG 301  
Db 395 GGAAGCGGCGTCCCGACACGACATACAGCGCCACGATGATGCTGTTGGGCGG 454  
QY 302 CTGCTCTCTGTTTCGCTATGTAGTGGGGGATCTCTCGGGATCTCTCTTCTGCTGCTCC 361  
Db 455 CTGCTCTCTGTTTCGCTATGTAGTGGGGGACCTCTCGGGATCTCTCTTCTGCTGCTCC 514  
QY 362 AGCTGTTTCAACATCTGCTGCTCGCGGATGAGAGCGGTGACGACTGCAATTGCTCAATCT 421  
Db 515 AGCTGTTTCAACATCTGCTGCTCGCGGATGAGAGCGGTGACGACTGCAATTGCTCAATCT 574  
QY 422 ATCCCGGCGACATTAACAGGTCAACCGGTTCAGGTATGGGATATGATGATGATGATGAT 481  
Db 575 ATCCCGGCGACATTAACCGGTTCAGGTATGGGATATGATGATGATGATGATGATGATGAT 634  
QY 482 CAACGCGGCTGCTGATCGCAGCTGCTCGGATCCCAAGCTGCTGCTGAGCATGCTGG 541  
Db 635 CAACGCGGCTGCTGATCGCAGCTGCTCGGATCCCAAGCTGCTGCTGAGCATGCTGG 694  
QY 542 CGGGGGCCATTGGGGAGTCTCGCGGCTCGCTGCTACTATTCCATGTTGGGAACTGG 601  
Db 695 CGGGGGCCATTGGGGAGTCTCGCGGCTCGCTGCTACTATTCCATGTTGGGAACTGG 754  
QY 602 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642  
Db 755 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795

## RESULT 4

US-08-927-597-5  
; Sequence 5, Application US/08927597  
; Patent No. 6245503  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,597  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/612,973  
; FILING DATE: 11-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:





```

Db 455 CTGCTTTCTGTTCCGCTATGTACGTGGGGACCTCTGGGATCTGCTTCTCTGCTCTCC 514
Qy 362 AGCTGTTTCAACATCTGCTCTGCTGGGATGAGACGGTGCAGGACTGCAATGCTCAATCT 421
Db 515 AGCTGTTTCAACATCTGCTCTGCTGGGATGAGACGGTGCAGGACTGCAATGCTCAATCT 574
Qy 422 ATCCCGGCCACATACAGGTCACCGTATGCTGGGATGATGATGATGATGATGATGATGAT 481
Db 575 ATCCCGGCCACATACAGGTCACCGTATGCTGGGATGATGATGATGATGATGATGATGAT 634
Qy 482 CAACGGCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Db 635 CAACGGCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 694
Qy 542 CGGGGGCCATGTTGGGAGTCTCTGGGGCCCTGCTACTATATCCATGTTGGGAACTGGG 601
Db 695 CGGGGGCCATGTTGGGAGTCTCTGGGGCCCTGCTACTATATCCATGTTGGGAACTGGG 754
Qy 602 CTAGGTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 632
Db 755 CTAGGTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 785

```

RESULT 6

```

US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079

```

```

; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
US-08-927-597-47

Query Match 96.5%; Score 619.8; DB 3; Length 2082;
Best Local Similarity 98.9%; Pred. No. 4.7e-155; Indels 0; Gaps 0;
Matches 624; Conservative 0; Mismatches 7;

Qy 2 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTTACTGCTCTGCTGACCAATCCAG 61
Db 155 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTTACTGCTCTGCTGACCAATCCAG 214
Qy 62 CTTCCGCTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTCTACGAAACGACTGCTCCA 121
Db 215 CTTCCGCTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTCTACGAAACGACTGCTCCA 274
Qy 122 ACTCAAGCAATTTGTTATGAGGCGGACATGATATGCACACCCCGGCTGCGTCCCT 181
Db 275 ACTCAAGCAATTTGTTATGAGGCGGACATGATATGCACACCCCGGCTGCGTCCCT 334
Qy 182 GCGTTCCGGGAGAACTCTTCTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTA 241
Db 335 GCGTTCCGGGAGAACTCTTCTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTA 394
Qy 242 GGAACGCCAGCGTCCCAACCAACATACGACGCCAGTGTGCTGTTGGGGCGG 301
Db 395 GGAACGCCAGCGTCCCAACCAACATACGACGCCAGTGTGCTGTTGGGGCGG 454
Qy 302 CTGCTCTCTGTTCCGCTATGTTAGTGGGGGATCTCTGGGGATCTCTCTCTGCTCTCCC 361
Db 455 CTGCTCTCTGTTCCGCTATGTTAGTGGGGGATCTCTGGGGATCTCTCTCTGCTCTCCC 514
Qy 362 AGCTGTTTCAACATCTGCTCTGCTGGGATGAGACGGTGCAGGACTGCAATGCTCAATCT 421
Db 515 AGCTGTTTCAACATCTGCTCTGCTGGGATGAGACGGTGCAGGACTGCAATGCTCAATCT 574
Qy 422 ATCCCGGCCACATACAGGTCACCGTATGCTGGGATGATGATGATGATGATGATGATGAT 481
Db 575 ATCCCGGCCACATACAGGTCACCGTATGCTGGGATGATGATGATGATGATGATGATGAT 634
Qy 482 CAACGGCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Db 635 CAACGGCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 694
Qy 542 CGGGGGCCATGTTGGGAGTCTCTGGGGCCCTGCTACTATATCCATGTTGGGAACTGGG 601
Db 695 CGGGGGCCATGTTGGGAGTCTCTGGGGCCCTGCTACTATATCCATGTTGGGAACTGGG 754
Qy 602 CTAGGTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 632
Db 755 CTAGGTTTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 785

```

RESULT 7

```

US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION NUMBER: US/08/612,973  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-10  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2433 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2430  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..2427  
US-08-612-973-49

Query Match 96.5%; Score 619.8; DB 3; Length 2433;  
Best Local Similarity 98.9%; Pred. No. 4.9e-155;  
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TGCCCGGTTGCTTTCTCTATCTTCTCTGTTGGTTTACTGTCTGTCTGACCAATTCAG 61  
Db |||||  
Qy 506 TGCCCGGTTGCTTTCTCTATCTTCTCTGTTGGTTTACTGTCTGTCTGACCAATTCAG 565  
Db |||||  
Qy 62 CTTCCGCTTATAGGTTGGGCAACGTTGTCGGGATGTACCAATGTACCAAGCACTGCCA 121  
Db |||||  
Qy 566 CTTCCGCTTATAGGTTGGGCAACGTTGTCGGGATGTACCAATGTACCAAGCACTGCCA 625  
Db |||||  
Qy 122 ACTCAAGCATTTGTATAGGAGCAGCGACATGATCATGCACACCCCGGGTGGCGCT 181  
Db |||||  
Qy 626 ACTCAAGCATTTGTATAGGAGCAGCGACATGATCATGCACACCCCGGGTGGCGCT 685  
Db |||||  
Qy 182 CGGTTCCGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGTA 241  
Db |||||  
Qy 686 CGGTTCCGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGTA 745  
Db |||||  
Qy 242 GGAACGCGAGCTCCCAACAGCAATACAGCCACGTCGATTTGCTGTTGGGGGG 301  
Db |||||  
Qy 746 GGAACGCGAGCTCCCAACAGCAATACAGCCACGTCGATTTGCTGTTGGGGGG 805  
Db |||||  
Qy 302 CTGCTCTCTGTTCCGCTATGATGATGCGGATCTCTCGGATCTGTTCTCTGCTCCC 361  
Db |||||  
Qy 806 CTGCTCTCTGTTCCGCTATGATGATGCGGATCTCTCGGATCTGTTCTCTGCTCCC 865  
Db |||||  
Qy 362 AGCTGTTTACCATCTCGCTCCCGGATGAGAGCGTGCAGGATGCAATTTGCTCAATCT 421  
Db |||||  
Qy 866 AGCTGTTTACCATCTCGCTCCCGGATGAGAGCGTGCAGGATGCAATTTGCTCAATCT 925  
Db |||||  
Qy 422 ATCCCGGCGCAATACAGGTCACCGTATGCTTGGATATGATGATGATGATGATGATGAT 481  
Db |||||  
Qy 926 ATCCCGGCGCAATACAGGTCACCGTATGCTTGGATATGATGATGATGATGATGATGAT 985  
Db |||||  
Qy 482 CAACGCGCTGTTGGTATGCGAGCTGCTCCGATCCCAAGCTGTCGTGGACATGTTGG 541  
Db |||||  
Qy 986 CAACGCGCTGTTGGTATGCGAGCTGCTCCGATCCCAAGCTGTCGTGGACATGTTGG 1045  
Db |||||

Qy 542 CGGGGGCCCATTTGGGGAGTCTCTGGCGGCTCTGGCTACTATTCTATCTGTTGGGAACTGGG 601  
Db |||||  
Qy 1046 CGGGGGCCCATTTGGGGAGTCTCTGGCGGCTCTGGCTACTATTCTATCTGTTGGGAACTGGG 1105  
Db |||||  
Qy 502 CTAAGGTTTGTATTTGATGTCATGCTACTCTTTTC 632  
Db |||||  
Qy 1106 CTAAGGTTTGTATTTGATGTCATGCTACTCTTTTC 1136  
Db |||||

## RESULT 8

US-08-927-597-49  
Sequence 49, Application US/08927597  
Patent No. 6245503  
GENERAL INFORMATION:  
APPLICANT: MAERTENS, GEERT  
APPLICANT: BOSMAN, FONS  
APPLICANT: DE MARTYNOFF, GUY  
APPLICANT: BUYSSE, MARIE-ANGE  
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION NUMBER: US/08/927,597  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/612,973  
FILING DATE: 11-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-10  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2433 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2430  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..2427  
US-08-927-597-49

Query Match 96.5%; Score 619.8; DB 3; Length 2433;  
Best Local Similarity 98.9%; Pred. No. 4.9e-155;  
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TGCCCGGTTGCTTTCTCTATCTTCTCTGTTGGTTTACTGTCTGTCTGACCAATTCAG 61  
Db |||||  
Qy 506 TGCCCGGTTGCTTTCTCTATCTTCTCTGTTGGTTTACTGTCTGTCTGACCAATTCAG 565  
Db |||||  
Qy 62 CTTCCGCTTATAGGTTGGGCAACGTTGTCGGGATGTACCAATGTACCAAGCACTGCCA 121  
Db |||||







; FILING DATE: 12-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weillacher, Robert G.  
; REGISTRATION NUMBER: 20,531  
; REFERENCE/DOCKET NUMBER: 06/59-47083.1  
; TELEPHONE: (202) 659-2811  
; TELEFAX: (202) 659-1462  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1539 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; US-08-470-426B-17

Query Match 87.8%; Score 563.8; DB 2; Length 1539;  
Best Local Similarity 93.3%; Pred. No. 3e-140;  
Matches 589; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTACTGCTGCTGACCAATCCAG 61  
DB |||||  
QY 506 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTGCTGCTGTTGACCATCCAG 565  
DB |||||  
QY 62 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGATACCATGTCAAGACGACTGCTCA 121  
DB |||||  
QY 566 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGATACCATGTCAAGACGACTGCTCA 625  
DB |||||  
QY 122 ACTCAAGCATTTGTTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGTGCCT 181  
DB |||||  
QY 626 ACTCAAGCATTTGTTATGAGCGAGCGACATGATCATGCACATCTCCCGGGTGGTGCCT 685  
DB |||||  
QY 182 GCGTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCACGCTGCGACGTA 241  
DB |||||  
QY 686 GCGTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCTCCACGCTGCGGCCA 745  
DB |||||  
QY 242 GGAACGCGAGCTCCCGACACAGCAATACGACGCGATGCGATGCGATGCTGTTGGGGCGG 301  
DB |||||  
QY 746 GGAATGCCAGGGTCCCGACATACGACGCGCAATACGACGCGCACTGCTGTTGGGGCGG 805  
DB |||||  
QY 302 CTGCTCTCTGCTCGCTATGATGAGCGGATCTCTCGGATGCTGCTGCTGCTGCC 361  
DB |||||  
QY 806 CTGCTCTCTGCTCGCTATGATGAGCGGATCTCTCGGATGCTGCTGCTGCTGCC 865  
DB |||||  
QY 362 AGCTGTTCAACATCTGCTCTGCGGATGAGAGCGGTGCGAGCTGCAATGCTCAATCT 421  
DB |||||  
QY 866 AGCTGTTCACTTCTGCTCTGCGGATGAGAGCGGTGCGAGCTGCAATGCTCAATCT 925  
DB |||||  
QY 422 ATCCCGGCGACATAAAGGTCACGTTATGCTGAGGATGCTGAGGATGATGATGAGTGGCCTA 481  
DB |||||  
QY 926 ATCCCGGCGACATAAAGGTCACGTTATGCTGAGGATGCTGAGGATGATGATGAGTGGCCTA 985  
DB |||||  
QY 482 CAACGCGCTGCTGCTGATGAGGATGCTGCTGAGTCCCGATCCCGATGCTGAGGATGCTGG 541  
DB |||||  
QY 986 CAACGCGCTGCTGCTGATGAGGATGCTGCTGAGTCCCGATCCCGATGCTGAGGATGCTGG 1045  
DB |||||  
QY 542 CGGGGGCCCATTTGGGAGTCTGCGGGCTCTGCGCTACTATTTCCATGCTGGGAACTGGG 601  
DB |||||  
QY 1046 CGGGGGCCCATTTGGGAGTCTGCGGGCTCTGCGCTACTATTTCCATGCTGGGAACTGGG 1105  
DB |||||  
QY 602 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632  
DB |||||  
QY 1106 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136  
DB |||||

## RESULT 14

US-08-470-426B-14  
; Sequence 14, Application US/08470426B  
; Patent No. 5856458  
; GENERAL INFORMATION:  
; APPLICANT: Okamoto, Hiroaki  
; APPLICANT: Nakamura, Tetsuo

; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR  
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B  
; TITLE OF INVENTION: HEPATITIS VIRUS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,  
; ADDRESSEE: L.L.P.  
; STREET: 1850 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,426B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-153402  
; FILING DATE: 12-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weillacher, Robert G.  
; REGISTRATION NUMBER: 20,531  
; REFERENCE/DOCKET NUMBER: 06/59-47083.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 659-2811  
; TELEFAX: (202) 659-1462  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1863 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; US-08-470-426B-14

Query Match 87.8%; Score 563.8; DB 2; Length 1863;  
Best Local Similarity 93.3%; Pred. No. 3.2e-140;  
Matches 589; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTACTGCTGCTGCTGACCAATCCAG 61  
DB |||||  
QY 830 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCT 889  
DB |||||  
QY 62 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGATACCATGTCAAGACGACTGCTCCA 121  
DB |||||  
QY 890 CTTCCGCTTATGAGTGGCGCAACGTGTCGGGATGATACCATGTCAAGACGACTGCTCCA 949  
DB |||||  
QY 122 ACTCAAGCATTTGTTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGTGCCT 181  
DB |||||  
QY 950 ACTCAAGCATTTGTTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGTGCCT 1009  
DB |||||  
QY 182 GCGTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGCTGCGAGCTA 241  
DB |||||  
QY 1010 GCGTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCTCCACGCTGCGGCCA 1069  
DB |||||  
QY 242 GGAACGCGAGCTCCCGACACAGCAATACGACGCGATGCGATGCTGCTGTTGGGGCGG 301  
DB |||||  
QY 1070 GGAATGCCAGGGTCCCGACATACGACGCGCAATACGACGCGCACTGCTGCTGTTGGGGCGG 1129  
DB |||||  
QY 302 CTGCTCTCTGTTCCGCTATGATGAGCGGATCTCTGCGGATGCTGCTGCTGCTGCTGCC 361  
DB |||||  
QY 1130 CTGCTCTCTGCTCGCTATGATGAGCGGATCTCTGCGGATGCTGCTGCTGCTGCC 1189  
DB |||||  
QY 362 AGCTGTTCAACATCTGCGCTGCTGCGGATGAGAGCGGTGCGAGCTGCAATGCTCAATCT 421  
DB |||||  
QY 1190 AGCTGTTCACTTCTGCGCTGCTGCGGATGAGAGCAAGTGCAGGACTGCAATCT 1249  
DB |||||  
QY 422 ATCCCGGCGACATAAAGGTCACGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481  
DB |||||





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:21:34 ; Search time 2231.49 Seconds  
(without alignments)  
10123.074 Million cell updates/sec

Title: US-09-899-303A-3  
Perfect score: 642  
Sequence: 1 ATGCCCGGTGCTCTTCTC.....TACTCTTGCTCTCTATATAG 642

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA Main:  
1: /cgn2\_6/ptodata/2/pna/US06 COMB.seq:  
2: /cgn2\_6/ptodata/2/pna/US06 COMB.seq:  
3: /cgn2\_6/ptodata/2/pna/US07 COMB.seq:  
4: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:  
5: /cgn2\_6/ptodata/2/pna/US081 COMB.seq:  
6: /cgn2\_6/ptodata/2/pna/US082 COMB.seq:  
7: /cgn2\_6/ptodata/2/pna/US083 COMB.seq:  
8: /cgn2\_6/ptodata/2/pna/US084 COMB.seq:  
9: /cgn2\_6/ptodata/2/pna/US085 COMB.seq:  
10: /cgn2\_6/ptodata/2/pna/US086 COMB.seq:  
11: /cgn2\_6/ptodata/2/pna/US087 COMB.seq:  
12: /cgn2\_6/ptodata/2/pna/US088 COMB.seq:  
13: /cgn2\_6/ptodata/2/pna/US089 COMB.seq:  
14: /cgn2\_6/ptodata/2/pna/US090 COMB.seq:  
15: /cgn2\_6/ptodata/2/pna/US091 COMB.seq:  
16: /cgn2\_6/ptodata/2/pna/US092 COMB.seq:  
17: /cgn2\_6/ptodata/2/pna/US093 COMB.seq:  
18: /cgn2\_6/ptodata/2/pna/US094 COMB.seq:  
19: /cgn2\_6/ptodata/2/pna/US095 COMB.seq:  
20: /cgn2\_6/ptodata/2/pna/US096 COMB.seq:  
21: /cgn2\_6/ptodata/2/pna/US097 COMB.seq:  
22: /cgn2\_6/ptodata/2/pna/US098 COMB.seq:  
23: /cgn2\_6/ptodata/2/pna/US099 COMB.seq:  
24: /cgn2\_6/ptodata/2/pna/US100 COMB.seq:  
25: /cgn2\_6/ptodata/2/pna/US101 COMB.seq:  
26: /cgn2\_6/ptodata/2/pna/US102 COMB.seq:  
27: /cgn2\_6/ptodata/2/pna/US103 COMB.seq:  
28: /cgn2\_6/ptodata/2/pna/US104 COMB.seq:  
29: /cgn2\_6/ptodata/2/pna/US105 COMB.seq:  
30: /cgn2\_6/ptodata/2/pna/US106 COMB.seq:  
31: /cgn2\_6/ptodata/2/pna/US107 COMB.seq:  
32: /cgn2\_6/ptodata/2/pna/US108 COMB.seq:  
33: /cgn2\_6/ptodata/2/pna/US109 COMB.seq:  
34: /cgn2\_6/ptodata/2/pna/US110 COMB.seq:  
35: /cgn2\_6/ptodata/2/pna/US111 COMB.seq:  
36: /cgn2\_6/ptodata/2/pna/US112 COMB.seq:  
37: /cgn2\_6/ptodata/2/pna/US113 COMB.seq:  
38: /cgn2\_6/ptodata/2/pna/US114 COMB.seq:  
39: /cgn2\_6/ptodata/2/pna/US115 COMB.seq:  
40: /cgn2\_6/ptodata/2/pna/US116 COMB.seq:  
41: /cgn2\_6/ptodata/2/pna/US117 COMB.seq:  
42: /cgn2\_6/ptodata/2/pna/US118 COMB.seq:  
43: /cgn2\_6/ptodata/2/pna/US119 COMB.seq:  
44: /cgn2\_6/ptodata/2/pna/US120 COMB.seq:  
45: /cgn2\_6/ptodata/2/pna/US121 COMB.seq:  
46: /cgn2\_6/ptodata/2/pna/US122 COMB.seq:  
47: /cgn2\_6/ptodata/2/pna/US123 COMB.seq:  
48: /cgn2\_6/ptodata/2/pna/US124 COMB.seq:  
49: /cgn2\_6/ptodata/2/pna/US125 COMB.seq:  
50: /cgn2\_6/ptodata/2/pna/US126 COMB.seq:  
51: /cgn2\_6/ptodata/2/pna/US127 COMB.seq:  
52: /cgn2\_6/ptodata/2/pna/US128 COMB.seq:  
53: /cgn2\_6/ptodata/2/pna/US129 COMB.seq:  
54: /cgn2\_6/ptodata/2/pna/US130 COMB.seq:  
55: /cgn2\_6/ptodata/2/pna/US131 COMB.seq:  
56: /cgn2\_6/ptodata/2/pna/US132 COMB.seq:  
57: /cgn2\_6/ptodata/2/pna/US133 COMB.seq:  
58: /cgn2\_6/ptodata/2/pna/US134 COMB.seq:  
59: /cgn2\_6/ptodata/2/pna/US135 COMB.seq:  
60: /cgn2\_6/ptodata/2/pna/US136 COMB.seq:  
61: /cgn2\_6/ptodata/2/pna/US137 COMB.seq:  
62: /cgn2\_6/ptodata/2/pna/US138 COMB.seq:  
63: /cgn2\_6/ptodata/2/pna/US139 COMB.seq:  
64: /cgn2\_6/ptodata/2/pna/US140 COMB.seq:  
65: /cgn2\_6/ptodata/2/pna/US141 COMB.seq:  
66: /cgn2\_6/ptodata/2/pna/US142 COMB.seq:  
67: /cgn2\_6/ptodata/2/pna/US143 COMB.seq:  
68: /cgn2\_6/ptodata/2/pna/US144 COMB.seq:  
69: /cgn2\_6/ptodata/2/pna/US145 COMB.seq:  
70: /cgn2\_6/ptodata/2/pna/US146 COMB.seq:  
71: /cgn2\_6/ptodata/2/pna/US147 COMB.seq:  
72: /cgn2\_6/ptodata/2/pna/US148 COMB.seq:  
73: /cgn2\_6/ptodata/2/pna/US149 COMB.seq:  
74: /cgn2\_6/ptodata/2/pna/US150 COMB.seq:

The  
Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The  
Pending databases have been split into two parts to reduce the amount of time required for their daily  
updates. This results in more machine time being available for processing searches, with the extensions  
Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions  
.rapm and .rapn  
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions  
.rapm and .rapn

Because they contain data that is confidential, the results of Pending  
database searches should not be left in the case.

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 2425.41 Seconds  
(without alignments)  
14206.949 Million cell updates/sec

Title: US-09-899-303a-5  
Perfect score: 795  
Sequence: 1 ATGTTGGTAAGTCAATGCA.....TACTCTTGTCTCCTAATAG 795

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vt.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	795	100.0	795	6	A48667	A48667 Sequence 5
2	795	100.0	795	6	AR157325	AR157325 Sequence
3	795	100.0	795	6	AX452754	AX452754 Sequence
4	795	100.0	795	6	AX685006	AX685006 Sequence
5	795	100.0	795	6	AX794848	AX794848 Sequence
6	778.8	98.0	2082	6	A48709	A48709 Sequence 47
7	778.8	98.0	2082	6	AR157350	AR157350 Sequence
8	778.8	98.0	2082	6	AX452796	AX452796 Sequence
9	778.8	98.0	2082	6	AX685048	AX685048 Sequence
10	778.8	98.0	2082	6	AX794890	AX794890 Sequence
11	778.8	98.0	2433	6	A48711	A48711 Sequence 49
12	778.8	98.0	2433	6	AR157351	AR157351 Sequence
13	778.8	98.0	2433	6	AX452798	AX452798 Sequence
14	778.8	98.0	2433	6	AX685050	AX685050 Sequence
15	778.8	98.0	2433	6	AX794892	AX794892 Sequence
16	711.6	89.5	9379	14	AF207766	AF207766 Hepatitis
17	711.6	89.5	9379	14	HCVPOLYP	AJ7000009 Hepatitis
18	711.6	89.5	9379	14	HCVPOLYP	AJ132996 Hepatitis
19	710	89.3	9386	14	AF165056	AF165056 Hepatitis
20	708.4	89.1	9386	14	AF165055	AF165055 Hepatitis
21	708.4	89.1	9410	14	HPCKIR2	D50481 Hepatitis C
22	706.8	88.9	3296	14	AB008441	AB008441 Hepatitis
23	706.8	88.9	3296	14	AB008443	AB008443 Hepatitis
24	706.8	88.9	9379	14	AF165052	AF165052 Hepatitis
25	706.8	88.9	9379	14	AF207761	AF207761 Hepatitis
26	705.2	88.7	3296	14	AB008442	AB008442 Hepatitis
27	705.2	88.7	3296	14	AB008446	AB008446 Hepatitis
28	705.2	88.7	9379	14	AF165051	AF165051 Hepatitis
29	705.2	88.7	9598	14	AB049101	AB049101 Hepatitis
30	703.6	88.5	3296	14	AB008447	AB008447 Hepatitis
31	703.6	88.5	9547	14	AB049091	AB049091 Hepatitis
32	703.6	88.5	9605	6	AX739971	AX739971 Sequence
33	703.6	88.5	9605	14	HCJ238799	AJ238799 Hepatitis
34	703.6	88.5	11076	6	AR406040	AR406040 Sequence
35	703.6	88.5	11076	6	AR406046	AR406046 Sequence
36	703.6	88.5	11076	6	AR406048	AR406048 Sequence
37	703.6	88.5	11076	6	AR406050	AR406050 Sequence
38	703.6	88.5	11076	6	AX036252	AX036252 Sequence
39	703.6	88.5	11076	6	AX036258	AX036258 Sequence
40	703.6	88.5	11076	6	AX036260	AX036260 Sequence
41	703.6	88.5	11076	6	AX036262	AX036262 Sequence
42	702	88.3	1880	14	HPC5TRJ4	D00832 Hepatitis C
43	702	88.3	2540	6	E04260	E04260 cDNA encodi
44	702	88.3	2540	6	E04805	E04805 cDNA to 5'-
45	702	88.3	2540	6	E07391	E07391 cDNA encodi

ALIGNMENTS

RESULT 1

LOCUS	A48667	795 bp	DNA	linear	PAT 07-MAR-1997
DEFINITION	Sequence 5 from Patent WO9604385.				
ACCESSION	A48667				
VERSION	A48667.1	GI:2302380			
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 795)				
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buyse,M.				
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE				
JOURNAL	Patent: WO 9604385-A 5 15-FEB-1996;				

```
INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
    source
        1..795
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
    CDS
        1..792
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAA03129.1"
            /db_xref="GI:2302381"
            /translation="MLGKVIDITLTCGFADLVGIPLVGAPLGGARALAHGVRVLEDD
            VNYATGMLPGCSFISIPLLALLSCLTPASAYEVNVSVMVHTNDCNSISIVYRAADM
            IMHTPCVPCVENSSRCVVALTPTLAARASVPTTIRRHVDLLVGAARFCSAMV
            GDLCGSFVLVSQIFTSIPRHETVQDCNSIYPGHITGRHMDMMWNSPTTALVVS
            QLLRIPQAVDVMVAGAHGVLGAGLAYTSMVGNWAKVLIYMLLPAP"
            mat_peptide
                1..789
                    /product="unnamed"
ORIGIN
    Query Match      100.0%; Score 795; DB 6; Length 795;
    Best Local Similarity 100.0%; Pred. No. 3e-167;
    Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCGGGTT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCGGGTT 120
QY 121 CTGGAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTGGGCTTTGCTGCTCTGCTGACCGCTTCAGAGCTTCGCGTTCAGAGTGCAGCGT 240
DB 181 CTCTGGGCTTTGCTGCTCTGCTGACCGCTTCAGAGCTTCGCGTTCAGAGTGCAGCGT 240
QY 241 TCGGGATGTACCATGTGTCAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGCGAG 300
DB 241 TCGGGATGTACCATGTGTCAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGCGAG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTACACCCCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 420
DB 361 TGCTGGGTAGCGCTACACCCCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 420
QY 421 ATACGAGCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGCTATGAGCGT 480
DB 421 ATACGAGCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGCTATGAGCGT 480
QY 481 GGGGACCTCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GGGGACCTCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 600
DB 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 600
QY 601 ATGGCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 ATGGCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CTCGGATCCCAAGCTGTGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 720
DB 661 CTCGGATCCCAAGCTGTGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 720

RESULT 2
AR157325
LOCUS AR157325 795 bp linear PAT 17-OCT-2001
DEFINITION Sequence 5 from parent US 6245503
ACCESSION AR157325
VERSION AR157325.1 GI:16218258
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
    1 (bases 1 to 795)
    Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.
    Purified hepatitis C virus envelope proteins for diagnostic and
    therapeutic use
    JOURNAL Patent: US 6245503-A 5 12-JUN-2001;
    LOCATION/Qualifiers
    source
        1..795
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
    Query Match      100.0%; Score 795; DB 6; Length 795;
    Best Local Similarity 100.0%; Pred. No. 3e-167;
    Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCGGGTT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCGGGTT 120
QY 121 CTGGAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTGGGCTTTGCTGCTCTGCTGACCGCTTCAGAGCTTCGCGTTCAGAGTGCAGCGT 240
DB 181 CTCTGGGCTTTGCTGCTCTGCTGACCGCTTCAGAGCTTCGCGTTCAGAGTGCAGCGT 240
QY 241 TCGGGATGTACCATGTGTCAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGCGAG 300
DB 241 TCGGGATGTACCATGTGTCAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGCGAG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTACACCCCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 420
DB 361 TGCTGGGTAGCGCTACACCCCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 420
QY 421 ATACGAGCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGCTATGAGCGT 480
DB 421 ATACGAGCGGCTGAGTGCCTGCGTTCGGGAGAACAACTCTTCCCGCTATGAGCGT 480
QY 481 GGGGACCTCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GGGGACCTCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 600
DB 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 600
QY 601 ATGGCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 ATGGCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CTCGGATCCCAAGCTGTGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 720
DB 661 CTCGGATCCCAAGCTGTGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCA 720
```



VNYATGNLPKGSFSLFLLALLSLCTVPASAYEVNRVSMHYVNDSCNSSIYVEAADM  
IMHTPGCVPCVRNNSSRCWVALTPTLAARNASVPTTIRRHVDLLVGAAPCSAMV  
GDLGCVFLVSLQFTTSPRHHVTQDCNCSLYPGHITGHRMADMMWNSPTTALVVS  
QLLRIPQAVDMVAGAHGVLAGLAYYSVNGWNAKVLIVMLLPAP  
1. .789  
/product="unnamed"

mat\_peptide

ORIGIN

Query Match 100.0%; Score 795; DB 6; Length 795;  
Best Local Similarity 100.0%; Pred. No. 3e-167;  
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60  
Db 1 ATGTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60  
Qy 61 CGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGT 120  
Db 61 CGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGT 120  
Qy 121 CTGGAGGACGGGTGAATCTATCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTC 180  
Db 121 CTGGAGGACGGGTGAATCTATCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTC 180  
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCCTGCTTATGAAGTGCACAACTG 240  
Db 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGCTTCAGCTTCCTGCTTATGAAGTGCACAACTG 240  
Qy 241 TCCGGGATGATACCTATGTCAGAAAGCTCTCAACTCAAGCATTTGTATGAGGACGG 300  
Db 241 TCCGGGATGATACCTATGTCAGAAAGCTCTCAACTCAAGCATTTGTATGAGGACGG 300  
Qy 301 GACATGATCATGCACACCCCGGCTGGTCCCTCGTTCGGGAGAAACAATCTTCCCGC 360  
Db 301 GACATGATCATGCACACCCCGGCTGGTCCCTCGTTCGGGAGAAACAATCTTCCCGC 360  
Qy 361 TGCTGGGTAGGCTCACCCCAAGCTCGCAGCTAGGAAACGCGAGCTCCCAACCAAGACA 420  
Db 361 TGCTGGGTAGGCTCACCCCAAGCTCGCAGCTAGGAAACGCGAGCTCCCAACCAAGACA 420  
Qy 421 ATAGAACGCGACGTCGATTTGCTTGGGGGGCGTCTTCTGTTCCGCTATGACGTG 480  
Db 421 ATAGAACGCGACGTCGATTTGCTTGGGGGGCGTCTTCTGTTCCGCTATGACGTG 480  
Qy 481 GGGGACCTCTGCGGATCTGCTTCTTCCCTCGCTCCAGCTGTTCACCATCTCGCTCCCG 540  
Db 481 GGGGACCTCTGCGGATCTGCTTCTTCCCTCGCTCCAGCTGTTCACCATCTCGCTCCCG 540  
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTCACT 600  
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTCACT 600  
Qy 601 ATGGCTTGGGATATGATGATGAATCTGCTCCCTACAGCGGCTCGTGGTATCGAGCTG 660  
Db 601 ATGGCTTGGGATATGATGATGAATCTGCTCCCTACAGCGGCTCGTGGTATCGAGCTG 660  
Qy 661 CTCGGGATCCCAAGCTGCTGTCGACATGTTGGGGGGCGCCATTTGGGGAGTCTTCGGG 720  
Db 661 CTCGGGATCCCAAGCTGCTGTCGACATGTTGGGGGGCGCCATTTGGGGAGTCTTCGGG 720  
Qy 721 GGTCTCGCTACTATTTCATGTTGGGGAATCTGGGCTAAGGTTTGAATGTTGATGCTACTC 780  
Db 721 GGTCTCGCTACTATTTCATGTTGGGGAATCTGGGCTAAGGTTTGAATGTTGATGCTACTC 780  
Qy 781 TTGTCTCCCTAATAG 795  
Db 781 TTGTCTCCCTAATAG 795

RESULT 5  
AX794848  
LOCUS  
DEFINITION Sequence 5 from Patent WO03051912.  
995 bp DNA linear PAT 04-OCT-2003

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AX794848  
AX794848.1 GI:37515737

Hepatitis C virus  
Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.

REFERENCE

1  
Maertens,G., Depla,E. and Boeman,F.  
Purified Hepatitis C virus envelope proteins for diagnostic and  
therapeutic use  
Patent: WO 03051912-A 5 26-JUN-2003;  
INNOGENETICS N.V. (BE)  
Location/Qualifiers  
1. .795  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
1. .792  
/note="unnamed protein product"

CDS

1. .789  
/product="unnamed"

mat\_peptide

ORIGIN

Query Match 100.0%; Score 795; DB 6; Length 795;  
Best Local Similarity 100.0%; Pred. No. 3e-167;  
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60  
Db 1 ATGTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60  
Qy 61 CGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGT 120  
Db 61 CGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGT 120  
Qy 121 CTGGAGGACGGGTGAATCTATCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTC 180  
Db 121 CTGGAGGACGGGTGAATCTATCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTC 180  
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCCTGCTTATGAAGTGCACAACTG 240  
Db 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCCTGCTTATGAAGTGCACAACTG 240  
Qy 241 TCCGGGATGATACCTATGTCAGAAAGCTCTCAACTCAAGCATTTGTATGAGGACGG 300  
Db 241 TCCGGGATGATACCTATGTCAGAAAGCTCTCAACTCAAGCATTTGTATGAGGACGG 300  
Qy 301 GACATGATCATGCACACCCCGGCTGGTCCCTCGTTCGGGAGAAACAATCTTCCCGC 360  
Db 301 GACATGATCATGCACACCCCGGCTGGTCCCTCGTTCGGGAGAAACAATCTTCCCGC 360  
Qy 361 TGCTGGGTAGGCTCACCCCAAGCTCGCAGCTAGGAAACGCGAGCTCCCAACCAAGACA 420  
Db 361 TGCTGGGTAGGCTCACCCCAAGCTCGCAGCTAGGAAACGCGAGCTCCCAACCAAGACA 420  
Qy 421 ATAGAACGCGACGTCGATTTGCTTGGGGGGCGTCTTCTGTTCCGCTATGACGTG 480  
Db 421 ATAGAACGCGACGTCGATTTGCTTGGGGGGCGTCTTCTGTTCCGCTATGACGTG 480  
Qy 481 GGGGACCTCTGCGGATCTGCTTCTTCCCTCGCTCCAGCTGTTCACCATCTCGCTCCCG 540  
Db 481 GGGGACCTCTGCGGATCTGCTTCTTCCCTCGCTCCAGCTGTTCACCATCTCGCTCCCG 540  
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTCACT 600





[illegible]



Db 484 GACCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTTCAACATCTCGCCTGCGCGCAT 543  
Qy 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTCACCGTATG 603  
Db 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTCACCGTATG 603  
Qy 604 GCTTGGGATATGATGATGAACTGGTTCGCTTACAAACGGCGCTGCTGATCGCAGCTGCTC 663  
Db 604 GCTTGGGATATGATGATGAACTGGTTCGCTTACAAACGGCGCTGCTGATCGCAGCTGCTC 663  
Qy 664 CGATCCCAACGCTGCTGCGACATGCTGCGCGCGCGCGCCATTTGGGGAGTCTTCTCTT 723  
Db 664 CGATCCCAACGCTGCTGCGACATGCTGCGCGCGCGCGCCATTTGGGGAGTCTTCTCTT 723  
Qy 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTTGGTGTGATGCTACTCTTT 783  
Db 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTTGGTGTGATGCTACTCTTT 783  
Qy 784 GC 785  
Db 784 GC 785

RESULT 9  
AX685048 2082 bp DNA linear PAT 29-MAR-2003  
LOCUS  
DEFINITION  
Sequence 47 from Patent WO02055548.  
ACCESSION  
AX685048  
VERSION  
AX685048.1 GI:29371453  
KEYWORDS  
Hepatitis C virus  
SOURCE  
Hepatitis C virus  
ORGANISM  
Hepatitis C virus  
REFERENCE  
1  
Maertens,G., Bosman,F. and Buyse,M.A.  
Purified Hepatitis C Virus envelope proteins for diagnostic and  
therapeutic use  
Patent: WO 02055548-A 47 18-JUL-2002;  
JOURNAL  
INNOGENETICS N.V. (BB)

FEATURES  
source  
1. .2082  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
<1. .2079  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD86535.1"  
/db\_xref="GI:29371454"  
/translation="NLTGKVIDTTCGPADLVGYIPLVGAFLGGAARALAHGVRVLEDG  
VYATGNLPCSFSLFLLALLSCLTPASAVEVNVSMVHVNDSCNSIVVEAADM  
IMTPGCPVCRVNNSSRCWALTPTLAARNAVPTTIRRHVDLLVGAARFSAWTV  
GDLGCVFLVSQFLTSPRHETVQDCNSYPCGHITGHRMAMDMNNWSPTRALVVS  
QLLRIPQAVDMVAGAHVGLAYYSVMGNNAKLVVWLLFAGVDGHTRVSGAAA  
SDTRGLVSPGSAQIQLVNTGNSWHINRTALNCNDSIQTFPAALFYKHFNSSG  
CPEPLASRIDKPAQGWGLTYTEPNSDQRPVCMHYAPRCPGIVPASPQVGVPCF  
TPSPVVTGTTDRGVPYNNWANDSVLIILNTRPPRGNMFGCTWNNGTGFTKTCGCP  
PCNIGGANNLTCPTDCPKHPEATYARCGSPWLTPRCMVHPYPLWHPCTVPT  
IFKRVMTGVGHRFDEAACNWTGERCDLDRKSEUSPLLLSTTEWQILPCSFITLP  
ALSTGITHQNTVDVQYDYGVSASVLIKWBYVLLFLLLADARICACLNWMLLI  
AQEAALENIVLNAAVAGAHGTLSPVFPFAAWYIKGRLVFGAAVAFYGVWPLLL  
LLALPPRAYA"  
mat\_peptide  
1. .2076  
/product="unnamed"

CDS  
1. .2079  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD86535.1"  
/db\_xref="GI:29371454"  
/translation="NLTGKVIDTTCGPADLVGYIPLVGAFLGGAARALAHGVRVLEDG  
VYATGNLPCSFSLFLLALLSCLTPASAVEVNVSMVHVNDSCNSIVVEAADM  
IMTPGCPVCRVNNSSRCWALTPTLAARNAVPTTIRRHVDLLVGAARFSAWTV  
GDLGCVFLVSQFLTSPRHETVQDCNSYPCGHITGHRMAMDMNNWSPTRALVVS  
QLLRIPQAVDMVAGAHVGLAYYSVMGNNAKLVVWLLFAGVDGHTRVSGAAA  
SDTRGLVSPGSAQIQLVNTGNSWHINRTALNCNDSIQTFPAALFYKHFNSSG  
CPEPLASRIDKPAQGWGLTYTEPNSDQRPVCMHYAPRCPGIVPASPQVGVPCF  
TPSPVVTGTTDRGVPYNNWANDSVLIILNTRPPRGNMFGCTWNNGTGFTKTCGCP  
PCNIGGANNLTCPTDCPKHPEATYARCGSPWLTPRCMVHPYPLWHPCTVPT  
IFKRVMTGVGHRFDEAACNWTGERCDLDRKSEUSPLLLSTTEWQILPCSFITLP  
ALSTGITHQNTVDVQYDYGVSASVLIKWBYVLLFLLLADARICACLNWMLLI  
AQEAALENIVLNAAVAGAHGTLSPVFPFAAWYIKGRLVFGAAVAFYGVWPLLL  
LLALPPRAYA"

mat\_peptide  
1. .2076  
/product="unnamed"

ORIGIN  
Query Match 98.0%; Score 778.8; DB 6; Length 2082;  
Best Local Similarity 99.7%; Pred. No. 1.2e-163;  
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

4 TTGGGTAAGTATCGATACCTTTATATGCGGCTTCGCGGACCTCGTGGGATACATTCG 63

Db 4 TTGGGTAAGTATCGATACCTTTATATGCGGCTTCGCGGACCTCGTGGGATACATTCG 63  
Qy 64 CTGCTGCGCGCCCTTCTAGGGGGGCTGCTGCGAGGGCCCTGCGCATGCGGCTCG 123  
Db 64 CTGCTGCGCGCCCTTCTAGGGGGGCTGCTGCGAGGGCCCTGCGCATGCGGCTCG 123  
Qy 124 GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCTC 183  
Db 124 GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCTC 183  
Qy 184 TTGGCTTTGCTGCTCTGCTGCAACCTTCCAGCTTCCGCTTATGAAAGTGGCAACGTGTC 243  
Db 184 TTGGCTTTGCTGCTCTGCTGCAACCTTCCAGCTTCCGCTTATGAAAGTGGCAACGTGTC 243  
Qy 244 GGGATGTACCATGTCAACAGCACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 303  
Db 244 GGGATGTACCATGTCAACAGCACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 303  
Qy 304 ATGATCATGCAACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
Db 304 ATGATCATGCAACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
Qy 364 TTGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
Db 364 TTGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
Qy 424 CGACCGCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
Db 424 CGACCGCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
Qy 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543  
Db 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543  
Qy 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTCACCGTATG 603  
Db 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTCACCGTATG 603  
Qy 604 GCTTGGGATATGATGATGAACTGGTTCGCTTCAACAGCGGCCCTGGTGGTATGCGAGCTGCTC 663  
Db 604 GCTTGGGATATGATGATGAACTGGTTCGCTTCAACAGCGGCCCTGGTGGTATGCGAGCTGCTC 663  
Qy 664 CGATCCCAACGCTGCTGCGACATGCTGCGGGGGGCCCATTTGGGGAGTCTTGGCGGGT 723  
Db 664 CGATCCCAACGCTGCTGCGACATGCTGCGGGGGGCCCATTTGGGGAGTCTTGGCGGGT 723  
Qy 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTTGGTGTGATGCTACTCTTT 783  
Db 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTTGGTGTGATGCTACTCTTT 783  
Qy 784 GC 785  
Db 784 GC 785

RESULT 10  
AX794890 2082 bp DNA linear PAT 04-OCT-2003  
LOCUS  
DEFINITION  
Sequence 47 from Patent WO03051912.  
ACCESSION  
AX794890  
VERSION  
AX794890.1 GI:37515779  
KEYWORDS  
Hepatitis C virus  
SOURCE  
Hepatitis C virus  
ORGANISM  
Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.

REFERENCE  
1  
Maertens,G., Depla,E. and Bosman,F.  
Purified Hepatitis C virus envelope proteins for diagnostic and  
therapeutic use  
Patent: WO 03051912-A 47 26-JUN-2003;  
JOURNAL  
INNOGENETICS N.V. (BB)

Db	604	GCTTGGGATATGATGATGAACACTGGTCGCCCTACACCGCCCTGGTGGTATCGCAGCTGCTC	663
Qy	664	CGGATCCCAAGCTGCTCGTGACATGCTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGGT	723
Db	664	CGGATCCCAAGCTGCTCGTGACATGCTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGGC	723
Qy	724	CTCGCCTACTATTCATGCTGGGGAACACTGGGCTTAAGCTTTTGATGTGTGACTACTCTTT	783
Db	724	CTCGCCTACTATTCATGCTGGGGAACACTGGGCTTAAGCTTTTGATGTGTGACTACTCTTT	783
Qy	784	GC 785	
Db	784	GC 785	
RESULT 11			
LOCUS	A48711	Sequence 49 from Patent WO9604385.	linear PAT 07-MAR-1997
DEFINITION	A48711	Sequence 49 from Patent WO9604385.	
ACCESSION	A48711.1	GI:2302424	
VERSION			
KEYWORDS			
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2433)		
AUTHORS	Maertens, G., Bosman, F., De M.G. and Buysse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	Patent: WO 9604385-A 49 15-FEB-1996;		
COMMENT	INNOGENETICS NV (BE)		
FEATURES	Other publication CA 2172273 960215		
	Other publication AU 3382495 960304.		
	Location/Qualifiers		
	1..2433		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
CDS	1..2430		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAA03147.1"		
	/db_xref="GI:2302425"		
	/db_xref="REMBL:CAA03147"		
	/translation="MSTNPKPQRKTRNTNRPPQDKVFPGGGQIVGGVYLLPRRGRPL GVATRAKTSRQPRROPIKARPEGRWAQPGYPMPLVGNMGWAGWLLSPRG SRPSWGTTPRRSRNLKVIDTLCGFADLVGYIPLVGPAGLGAARALAHGVRLV GVNYATGNLPFGCSIFLLALLSLVTPASAYEVNRVSCMYHVTNDCNSSIIVYBAAD MINHTPGCVCRVNRSSRCWALVTPLAARNASVPTTIRRHVDLLVGAAPCSAMY GDLGCVFLVSQLFTISPRRHETVQDCNCSYPGHI TGRHMAWMMWNSPTTALVVS QLLRIPQAVDMVAGHVLGALAYTSMVGNWAKLVVWMLLPAGVDCHTVSGCAA SDTRGLVSLFSPGSAQKIQLVNTGSHINFTALNCNDSLOTGFPFALPYKHKNSSG CPEPLASCRSIDKPAQGWGPLYTTEPNSSDQRPYCMHYAPRPGIIVPASQVGPVYCF TPSPVVGTTDRFQVTPYNGANDSDVLINANTPRGNWFGCTWMNGTGPTKCGGP PCNTGGAGNLTCTPTDCFRKHPEATYARCSGPMWTPRCMVHYPIRLMHPYCTVNF IPKVMYGVGVEHFEACNWRGRCDEDRSELSPLLSITTEWOLIPCSFTTLP ALSTGLIHLHONIVDVQVLYGVSAVSLVTKWEYVLLFLILLADARICACIMWMLL AQARALENLVLRNAAVAGAHGTLFLVFFCAAWYIKGRIVPGNAYAFYGVWPLLL LLALUPPPRAYA"		
	mat_peptide	1..2427	/product="unnamed"
ORIGIN			
Query Match 98.0%; Score 778.8; DB 6; Length 2433;			
Best Local Similarity 99.7%; Pred. No. 1.2e-163;			
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	4	TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG	63
Db	355	TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG	414
Qy	64	CTCGTCGGCGCCCCCTAGAGGGGCGCTGCGAGGGCCCTGGCGCATCGCGTTCG	123

Db 415 CTGCTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGCGATGGGGTCCGGGTTCTG 474  
Qy 124 GAGAGCGCGTGAACATATGCAACAGGGAAATTTGGCCGGTGTCTCTTTCTATCTTCTC 183  
Db 475 GAGAGCGCGTGAACATATGCAACAGGGAAATTTGGCCGGTGTCTCTTTCTATCTTCTC 534  
Qy 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 243  
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 594  
Qy 244 GCGATGTACCATGTGACAGACGATGCTTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 303  
Db 595 GCGATGTACCATGTGACAGACGATGCTTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 654  
Qy 304 ATGATCATGACACACCCCGGGTGTGCTGCTTCCGCTTCCGAGAGAACAACTTCCCGCTGC 363  
Db 655 ATGATCATGACACACCCCGGGTGTGCTGCTTCCGCTTCCGAGAGAACAACTTCCCGCTGC 714  
Qy 364 TGGGTAGCGCTCACCCCGGCTGCTGAGAGTGGAAACGCGAGCGTCCCGACCAACAATA 423  
Db 715 TGGGTAGCGCTCACCCCGGCTGCTGAGAGTGGAAACGCGAGCGTCCCGACCAACAATA 774  
Qy 424 CGAGCGACAGTGGCTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGCTGGG 483  
Db 775 CGAGCGACAGTGGCTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGCTGGG 834  
Qy 484 GACCTCTGCGGATCTGTCTTCCCTGCTCCAGCTGTTCAACCATCTCGCCCTGCGCGCAT 543  
Db 835 GACCTCTGCGGATCTGTCTTCCCTGCTCCAGCTGTTCAACCATCTCGCCCTGCGCGCAT 894  
Qy 544 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGTATG 603  
Db 895 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGTATG 954  
Qy 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
Db 955 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014  
Qy 784 GC 785  
Db 1135 GC 1136

RESULT 12  
LOCUS AR157351 2433 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 49 from patent US 6245503.  
ACCESSION AR157351  
VERSION AR157351.1 GI:16218285  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2433)  
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.  
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use  
JOURNAL Patent: US 6245503-A 49 12-JUN-2001;  
FEATURES Location/Qualifiers  
source 1..2433  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 98.0%; Score 778.8; DB 6; Length 2433;

Best Local Similarity 99.7%; Pred. No. 1.2e-163;  
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 4 TTGGGTAAAGGTCAATGATACCTTACATGCGCTTTCGCGACCTCGTGGGGTACATTCGG 63  
Db 355 TTGGGTAAAGGTCAATGATACCTTACATGCGCTTTCGCGACCTCGTGGGGTACATTCGG 414  
Qy 64 CTGCTGGGCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGGCGATGGGGTTCG 123  
Db 415 CTGCTGGGCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGGCGATGGGGTTCG 474  
Qy 124 GAGAGCGCGTGAACATATGCAACAGGGAAATTTGGCCGGTGTCTCTTTCTATCTTCTC 183  
Db 475 GAGAGCGCGTGAACATATGCAACAGGGAAATTTGGCCGGTGTCTCTTTCTATCTTCTC 534  
Qy 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 243  
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 594  
Qy 244 GCGATGTACCATGTGACAGACGATGCTTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 303  
Db 595 GCGATGTACCATGTGACAGACGATGCTTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 654  
Qy 304 ATGATCATGACACACCCCGGGTGTGCTGCTTCCGCTTCCGAGAGAACAACTTCCCGCTGC 363  
Db 655 ATGATCATGACACACCCCGGGTGTGCTGCTTCCGCTTCCGAGAGAACAACTTCCCGCTGC 714  
Qy 364 TGGGTAGCGCTCACCCCGGCTGCTGAGAGTGGAAACGCGAGCGTCCCGACCAACAATA 423  
Db 715 TGGGTAGCGCTCACCCCGGCTGCTGAGAGTGGAAACGCGAGCGTCCCGACCAACAATA 774  
Qy 424 CGAGCGACAGTGGCTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGCTGGG 483  
Db 775 CGAGCGACAGTGGCTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGCTGGG 834  
Qy 484 GACCTCTGCGGATCTGTCTTCCCTGCTCCAGCTGTTCAACCATCTCGCCCTGCGCGCAT 543  
Db 835 GACCTCTGCGGATCTGTCTTCCCTGCTCCAGCTGTTCAACCATCTCGCCCTGCGCGCAT 894  
Qy 544 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGTATG 603  
Db 895 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGTATG 954  
Qy 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
Db 955 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014  
Qy 664 CGGATCCCAAGCTGTGCTGGAATGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723  
Db 1015 CGGATCCCAAGCTGTGCTGGAATGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1074  
Qy 724 CTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTGATGATGATGATGATGATG 783  
Db 1075 CTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTGATGATGATGATGATG 1134  
Qy 784 GC 785  
Db 1135 GC 1136

RESULT 13  
LOCUS AX452798 2433 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 49 from Patent EP1211315.  
ACCESSION AX452798  
VERSION AX452798.1 GI:21712483  
KEYWORDS  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
REFERENCE 1  
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.-A.

TITLE Recombinant vectors for producing hcv envelope proteins  
JOURNAL Patent: EP 1211315-A 49 05-JUN-2002;  
Innogenetics N.V. (BE)

FEATURES  
source Location/Qualifiers

1..2433  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
CDS  
1..2430  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD38083.1"  
/db\_xref="GI:21712484"  
/translation="MSTNPKPQKTRNTNRRPDQVKFGGQIVGVVLLPRRGPRL  
GVRATKTSERSOPRRROPIPKARRPGRWAQPGYPWPLVNGMGWAGWLLSPRG  
SRPSWGTDPRESRNIGKVIDTLTCGFADLVGYIPLVAGPLGGARALAHGVRLSD  
GVNATGNLPGCSFSLIALLSLCTVPASAYEVNRVSGMYHTNDCSSIVYRAAD  
MIMHTPGCVPCVRENSSRCWALVTPLAARNASVPTTIRRHVDLLVGAAPFCSAMY  
VGDLCGSLVLSQFTISPRRHETVQDCNSIYPGHI TGRMAWMMNWSPTLAV  
SOLLRIPOAVDMVAGAHVGLAGLAYSMVGNWAKVLMVLLPAGVGHTRVSGAA  
ASDTRGLVSLSPGSAOKTQLVNTGSMHINRTALNCNDSLOTGPPAALPYKHKNSS  
GCPERLASCRSIDKPAQHGPLTYTEPNSSDQRPYCNHVAAPRPGIIVPASVQGPVYC  
FTSPVVGTTDRFGPTTNWAGNDSVLLNTRPPRGNWFGCTWNGGTGFKTCGG  
PPCNTGGAGNNTLCTCDCKRHPATYACGSGPWLTPRCMVHYPTLWHPYCTVNF  
TIPKRVMYGVGHEHFEACNWRGRCDEDRSELSPLLSLTEMILPCSTFTL  
PALSTGLIHLHONI VDVQVLYGVGSAVSLVIKWEVLLFLILLADARIACILMMLL  
IAQARAALENLVLNAAAVAGAHGTLFLVFFCAAWYIKGRLVPGAAAYAFYGVWPLLL  
LILLALPPRAYA"  
1..2427  
/product="unnamed"

mat\_peptide

ORIGIN

Query Match 98.0%; Score 778.8; DB 6; Length 2433;  
Best Local Similarity 99.7%; Pred. No. 1.2e-163;  
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 63  
DB 355 TTGGGTAAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 414

QY 64 CTGCTGGGCGCCCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 123  
DB 415 CTGCTGGGCGCCCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 474

QY 124 GAGGACGGCGTGAATATGCAAGGAAATTTGCCCGGTGCTTCTCTATCTTCCTC 183  
DB 475 GAGGACGGCGTGAATATGCAAGGAAATTTGCCCGGTGCTTCTCTATCTTCCTC 534

QY 184 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTTCAGCTTTCAGAGTGCACACGTGTC 243  
DB 535 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTTCAGCTTTCAGAGTGCACACGTGTC 594

QY 244 GGGATGACATGTCAGAAAGCTGCTCAACTCAAGCAATGCTGATGAGGACGCGGAC 303  
DB 595 GGGATGACATGTCAGAAAGCTGCTCAACTCAAGCAATGCTGATGAGGACGCGGAC 654

QY 304 ATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
DB 655 ATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714

QY 364 TGGGTAGCGCTACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
DB 715 TGGGTAGCGCTACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774

QY 424 CGACGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
DB 775 CGACGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834

QY 484 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
DB 835 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594

QY 544 GAGACGGTGCAGAGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCCGTATG 603  
DB 895 GAGACGGTGCAGAGCTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCCGTATG 954

QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
DB 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014

QY 664 CGGATCCCAACAAGCTGTGCTGGACATGCTGGGGGGGGCCCATTTGGGGAGTCTCTGGGGGT 723  
DB 1015 CGGATCCCAACAAGCTGTGCTGGACATGCTGGGGGGGGCCCATTTGGGGAGTCTCTGGGGGT 1074

QY 724 CTGCGCTACTATTTCCATGCTGGGAGACTGCGGCTTAAGGTTTTTGATGCTGCTACTCTTT 783  
DB 1075 CTGCGCTACTATTTCCATGCTGGGAGACTGCGGCTTAAGGTTTTTGATGCTGCTACTCTTT 1134

QY 784 GC 785  
DB 1135 GC 1136

RESULT 14  
AX685050 LOCUS 2433 bp DNA linear PAT 29-MAR-2003  
DEFINITION Sequence 49 from Patent WO0205548.  
ACCESSION AX685050  
VERSION AX685050.1 GI:29371455  
KEYWORDS  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Virus; sRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1  
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.  
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use  
JOURNAL Patent: WO 0205548-A 49 18-JUL-2002;  
INNOCENTICS N.V. (BE)  
FEATURES  
source Location/Qualifiers  
1..2433  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
CDS  
1..2430  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD86536.1"  
/db\_xref="GI:29371456"  
/translation="MSTNPKPQKTRNTNRRPDQVKFGGQIVGVVLLPRRGPRL  
GVRATKTSERSOPRRROPIPKARRPGRWAQPGYPWPLVNGMGWAGWLLSPRG  
SRPSWGTDPRESRNIGKVIDTLTCGFADLVGYIPLVAGPLGGARALAHGVRLSD  
GVNATGNLPGCSFSLIALLSLCTVPASAYEVNRVSGMYHTNDCSSIVYRAAD  
MIMHTPGCVPCVRENSSRCWALVTPLAARNASVPTTIRRHVDLLVGAAPFCSAMY  
VGDLCGSLVLSQFTISPRRHETVQDCNSIYPGHI TGRMAWMMNWSPTLAV  
SOLLRIPOAVDMVAGAHVGLAGLAYSMVGNWAKVLMVLLPAGVGHTRVSGAA  
ASDTRGLVSLSPGSAOKTQLVNTGSMHINRTALNCNDSLOTGPPAALPYKHKNSS  
GCPERLASCRSIDKPAQHGPLTYTEPNSSDQRPYCNHVAAPRPGIIVPASVQGPVYC  
FTSPVVGTTDRFGPTTNWAGNDSVLLNTRPPRGNWFGCTWNGGTGFKTCGG  
PPCNTGGAGNNTLCTCDCKRHPATYACGSGPWLTPRCMVHYPTLWHPYCTVNF  
TIPKRVMYGVGHEHFEACNWRGRCDEDRSELSPLLSLTEMILPCSTFTL  
PALSTGLIHLHONI VDVQVLYGVGSAVSLVIKWEVLLFLILLADARIACILMMLL  
IAQARAALENLVLNAAAVAGAHGTLFLVFFCAAWYIKGRLVPGAAAYAFYGVWPLLL  
LILLALPPRAYA"  
1..2427  
/product="unnamed"

mat\_peptide

ORIGIN

Query Match 98.0%; Score 778.8; DB 6; Length 2433;  
Best Local Similarity 99.7%; Pred. No. 1.2e-163;  
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 63  
DB 355 TTGGGTAAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 414

QY 64 CTGCTGGGCGCCCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 123  
DB 415 CTGCTGGGCGCCCTTACATCGGCTTCGCGACCTCGTGGGATACATTCGG 474

QY 124 GAGGACGGCGTGAATATGCAAGGAAATTTGCCCGGTGCTTCTCTATCTTCCTC 183  
DB 475 GAGGACGGCGTGAATATGCAAGGAAATTTGCCCGGTGCTTCTCTATCTTCCTC 534

QY 184 TTGGCTTTGCTGCTGCTGACCGTTTCAGCTTTCAGCTTTCAGAGTGCACACGTGTC 243  
DB 535 TTGGCTTTGCTGCTGCTGACCGTTTCAGCTTTCAGCTTTCAGAGTGCACACGTGTC 594

QY 244 GGGATGACATGTCAGAAAGCTGCTCAACTCAAGCAATGCTGATGAGGACGCGGAC 303  
DB 595 GGGATGACATGTCAGAAAGCTGCTCAACTCAAGCAATGCTGATGAGGACGCGGAC 654

QY 304 ATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
DB 655 ATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714

QY 364 TGGGTAGCGCTACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
DB 715 TGGGTAGCGCTACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774

QY 424 CGACGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
DB 775 CGACGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834

QY 484 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
DB 835 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594

Db	355		TTGGGTAAAGTCATCGATACCTTTACATGCGGCTTCGCGACCTCTGTGGGTACATCCG	414
Qy	64		CTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGTCCGGTTCG	123
Db	415		CTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGTCCGGTTCG	474
Qy	124		GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTCTATCTTCTC	183
Db	475		GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTCTATCTTCTC	534
Qy	184		TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTCCGCTTATGAGTGGCAAGTGTCC	243
Db	535		TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTCCGCTTATGAGTGGCAAGTGTCC	594
Qy	244		GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC	303
Db	595		GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC	654
Qy	304		ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCGAGAACAACTCTTCCCGTGC	363
Db	655		ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCGAGAACAACTCTTCCCGTGC	714
Qy	364		TGGGTAGCGCTCACCCCGGCTGCGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG	423
Db	715		TGGGTAGCGCTCACCCCGGCTGCGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG	774
Qy	424		CGACCGCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG	483
Db	775		CGACCGCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG	834
Qy	484		GACCTCTGCGGATCTGTTCTCTCGTCTCCAGCTGTTTCACTCTCCGCTGCGCGCAT	543
Db	835		GACCTCTGCGGATCTGTTCTCTCGTCTCCAGCTGTTTCACTCTCCGCTGCGCGCAT	894
Qy	544		GAGCGGTGACGAGTCTCAATGCTCAATCTATCCCGGCGCACATAAGCGGTTCACCGTATG	603
Db	895		GAGCGGTGACGAGTCTCAATGCTCAATCTATCCCGGCGCACATAAGCGGTTCACCGTATG	954
Qy	604		GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	663
Db	955		GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1014
Qy	664		CGGATCCACAAAGCTGCTGCGGACATGCTGGCGGGCGCCATTTGGGAGTCTCTGCGGGT	723
Db	1015		CGGATCCACAAAGCTGCTGCGGACATGCTGGCGGGCGCCATTTGGGAGTCTCTGCGGGT	1074
Qy	724		CTCGCTTACTTATCCATGGTGGGAACTGGGCTTAAGGTTTGTATGATGCTACTCTTT	783
Db	1075		CTCGCTTACTTATCCATGGTGGGAACTGGGCTTAAGGTTTGTATGATGCTACTCTTT	1134
Qy	784	GC	785	
Db	1135	GC	1136	

RESULT 15					
AX794892	LOCUS	2433 bp	DNA	linear	PAT 04-OCT-2003
AX794892	DEFINITION	Sequence 49 from Patent WO03051912.			
AX794892	ACCESSION				
AX794892.1	VERSION	GI:37515781			
Hepatitis C virus	SOURCE				
Hepatitis C virus	ORGANISM				
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.					
1	REFERENCE				
Maertens,G., Depla,B. and Bosman,P.	AUTHORS				
Purified Hepatitis C virus envelope proteins for diagnostic and therapeutic use	TITLE				
Patent: WO 03051912-A 49 26-JUN-2003;	JOURNAL				
INOGENETICS N.V. (BE)					

FEATURES	Location/Qualifiers	
source	1. .2433	
	/organism="Hepatitis C virus"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:11103"	
CDS	1. .2430	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAB48259.1"	
	/db_xref="GI:37515782"	
	/translation="MSTNPKPQRTKENTNRRPQDVKPPGGGQIVGGVLLPRRGPRLL	
	GVRATKTSRSQPRGRQPIKARRPGRWAQPGYPMPLYGNEGNGWAGWLLSPRG	
	SRPSWGPTDPRRSRNGKVIDTLCGFADLVGYIPLVGAPLGGARALAHGRVLED	
	GVNATGNLPCGSPFSIFALLLSCLTTPASAYEVNRVSGWYVNDNSNIVIVBAAD	
	MIHTPCGVCPCRENSSRQWALPTPLAARNASVPTTIRRHVDLLVGAAPFCSAMT	
	SQDLGCVFLVSOLFPIPRRHETVQDCNSIYPGHIYGRHMMWMMNWNWSPPTALVV	
	SGLLRIPQAVVDMVAGAHGVLGCLAYYSVGNWAKVLVLMFAGVDTGHTRVSGAA	
	ASDTRGLVSLPSPGSAQKIQLVNTGSHWNRNALNCNDSLTQTFPALPYKHKNSS	
	GCPLRLASCSIDKFAQWGLPTTYTPNSSDQRFYCMHYPAPQCGIYPASQVCGPVYC	
	FTSPVVGTTDFRGVPTYNWANDSVLLINLNRPPRGNWPGCTMNGTGTFTKTCGG	
	PCNIGAGNNLTCTPTDCPRKHPEATVARGSGPMLTPRCMVHYPYRLWHVPTVNF	
	TIPIKRVYGVGVHRFPAACNWTGRBCDLREDRSELSPLLSTTRWQIILPCSPPLL	
	PALSTGLIHLHONTVDYOYLGVGSAVUSLVIKWEYVLLPLLLADARIACIWNMLL	
	IAQRAALENLIVLNAARVAGAHGTLFLVFFCAAWYIKGRDVPQANATYFTGWPELLL	
	LLALAPPRAVA"	
	1. .2427	
	/product="unnamed"	
mat_peptide		
ORIGIN		
Query Match	98.0%; Score 778.8; DB 6; Length 2433;	
Best Local Similarity	99.7%; Pred. No. 1.2e-163;	
Matches 780; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Qy	4	TTGGGTAAAGTCATCGATACCTTTACATGCGGCTTCGCGACCTCTGTGGGTACATCCG 63
Db	355	TTGGGTAAAGTCATCGATACCTTTACATGCGGCTTCGCGACCTCTGTGGGTACATCCG 414
Qy	64	CTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGGTTCG 123
Db	415	CTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGGTTCG 474
Qy	124	GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTCTATCTTCTC 183
Db	475	GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTCTATCTTCTC 534
Qy	184	TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCC 243
Db	535	TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCC 594
Qy	244	GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC 303
Db	595	GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC 654
Qy	304	ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCGAGAACAACTCTTCCCGTGC 363
Db	655	ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCGAGAACAACTCTTCCCGTGC 714
Qy	364	TGGGTAGCGCTCACCCCGGCTGCGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 423
Db	715	TGGGTAGCGCTCACCCCGGCTGCGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 774
Qy	424	CGACCGCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db	775	CGACCGCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834
Qy	484	GACCTCTGCGGATCTGTTCTCTCGTCTCCAGCTGTTTCACTCTCCGCTGCGCGCAT 543
Db	835	GACCTCTGCGGATCTGTTCTCTCGTCTCCAGCTGTTTCACTCTCCGCTGCGCGCAT 894
Qy	544	GAGCGGTGACGAGTCTCAATGCTCAATCTATCCCGGCGCACATAAGCGGTTCACCGTATG 603
Db	895	GAGCGGTGACGAGTCTCAATGCTCAATCTATCCCGGCGCACATAAGCGGTTCACCGTATG 954

```

QY 604 GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC 663
Db |||||
QY 955 GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC 1014
Db |||||
QY 664 CGGATCCCAACAGCTCTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTCTGGCGGGT 723
Db |||||
QY 1015 CGGATCCCAACAGCTCTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTCTGGCGGGC 1074
Db |||||
QY 724 CTCGCCCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTTGATTTGATGCTACTCTTTT 783
Db |||||
QY 1075 CTCGCCCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTTGATTTGATGCTACTCTTTT 1134
Db |||||
QY 784 GC 785
Db ||
1135 GC 1136

```

Search completed: June 16, 2004, 11:39:16  
Job time : 2427.41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 299.119 Seconds  
(without alignments)  
11290.892 Million cell updates/sec

Title: US-09-899-303a-5

Perfect score: 795

Sequence: 1 ATGTTGGTAAGGTCAATCGA.....TACTCTTGTGCTCCTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	795	100.0	795	2	AAT12705 HCV E1 CO
2	795	100.0	795	6	AAL48914 Hepatitis
3	795	100.0	795	9	ADD55513 Hepatitis
4	778.8	98.0	2082	6	AAL48939 Hepatitis
5	778.8	98.0	2082	9	ADD55555 Hepatitis
6	778.8	98.0	2086	2	AAT12973 HCV E1 CO
7	778.8	98.0	2433	2	AAT12974 HCV E1 CO
8	778.8	98.0	2433	9	ADD55557 Hepatitis
9	767.8	96.6	2434	6	AAL48940 Hepatitis
10	703.6	88.5	9605	6	ABK91431 Hepatitis
11	703.6	88.5	9605	6	ABK91424 Hepatitis
12	703.6	88.5	9605	6	ABK91429 Hepatitis
13	703.6	88.5	9605	6	ABK91432 Hepatitis
14	703.6	88.5	9605	6	ABK91411 Hepatitis
15	703.6	88.5	9605	6	ABK91430 Hepatitis
16	703.6	88.5	9605	6	ABK91428 Hepatitis
17	703.6	88.5	9605	6	ABK91425 Hepatitis
18	703.6	88.5	9605	6	ABK91426 Hepatitis
19	703.6	88.5	9605	6	ABK91433 Hepatitis
20	703.6	88.5	9605	6	AAD25332 Hepatitis
21	703.6	88.5	9608	6	ABK91427 Hepatitis
22	703.6	88.5	11062	6	AAD25331 Hepatitis
23	703.6	88.5	11076	3	AAA98965 Hepatitis

## ALIGNMENTS

### RESULT 1

AAT12705  
ID AAT12705 standard; DNA; 795 BP.

XX AC AAT12705;

DT 23-SEP-1996 (first entry)

XX DB HCV E1 construct HCC110A.

KW HCV; E1; E2: disulphide bond cleavage; envelope protein; vaccine; human;  
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;  
KW ss.

XX OS Hepatitis C virus.

XX PN MO9604385-A2.

XX PD 15-FEB-1996.

XX PF 31-JUL-1995; 95WO-EP003031.

XX PR 29-JUL-1994; 94EP-00870132.

XX PA (INNO-) INNOGENETICS NV.

XX PI Maertens G, Bosman F, De Martynoff G, Buysse M;

XX DR WPI; 1996-129401/13.

XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins  
PT - in presence of di-sulphide bond cleavage agent, to produce proteins  
PT suitable for direct use in vaccines or diagnostic assays of HCV.

XX PS Claim 23; Fig 21; 146pp; English.

XX CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1  
CC and E2 protein coding sequence constructs. These sequences are included  
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.  
CC The recombinant proteins can then be isolated using a method of the  
CC invention. In the method, the envelope proteins are purified by carrying  
CC out a disulphide bond cleavage, or a reduction step with a disulphide  
CC bond cleavage agent, after lysis of recombinant host cells. The  
CC constructs containing the purified HCV envelope proteins can be used for  
CC vaccinating humans against HCV, for in vitro detection of HCV antibodies  
CC in a sample, and in a serotyping assay for detecting one or more

ABA03491 Cuticle p  
Aa433889 NANB hepa  
Aa633753 NANBHV ge  
Aa24467 NANB hepa  
Aa64068 Non-A, no  
Aa30386 5'UTR/COR  
Aa60672 Fragment  
Aa155222 Plasmid p  
Aa29628 Hepatitis  
Aa60673 Fragment  
Aa26981 HCV gene  
Aa03677 Hepatitis  
Aa80498 DNA encod  
Aa81559 Hepatitis  
Aa03960 Partial H  
Aa25517 Hepatitis  
Aa153723 Hepatitis  
Aa49655 Hepatitis  
Aa15363 Fragment  
Aa12242 Encodes P  
Aa20923 C10-E12 D  
Aa33038 HCV-S1 fu



serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques

Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 U; 0 Other;

Query Match 100.0%; Score 795; DB 2; Length 795;  
Best Local Similarity 100.0%; Pred. NO. 4.1e-204;  
Matches 795; Conservative 0; Mismatches 0; Indels

Qy	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCCGACGTCGTGGGGGTACATT	60
Db	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCCGACGTCGTGGGGGTACATT	60
Qy	61	CCGCTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCGCGGTT	120
Db	61	CCGCTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCGCGGTT	120
Qy	121	CTGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC	180
Db	121	CTGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTTCAGCTTTCGCTTATGAGTGGCGAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTTCAGCTTTCGCTTATGAGTGGCGAACGTG	240
Qy	241	TCCGGGATGTACCATGTTCACGAAACGACTGCTCCAACTCAAGCATTTGTATCAGGCAGCG	300
Db	241	TCCGGGATGTACCATGTTCACGAAACGACTGCTCCAACTCAAGCATTTGTATCAGGCAGCG	300
Qy	301	GACATGATCATGCACACCCCGGGTGCGTGCTTCCGGAAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGCACACCCCGGGTGCGTGCTTCCGGAAGAACAACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGACGCCAGCGTCCCAACCAACGACA	420
Db	361	TGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGACGCCAGCGTCCCAACCAACGACA	420
Qy	421	ATACGACGCCACGTGCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG	480
Db	421	ATACGACGCCACGTGCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG	480
Qy	481	GGGGACCTCTCGGGATCTGTCTTCTGCTCTCCAGCTGTTCACCATCTCGCCTGCGCCGG	540
Db	481	GGGGACCTCTCGGGATCTGTCTTCTGCTCTCCAGCTGTTCACCATCTCGCCTGCGCCGG	540
Qy	541	CATCAGACGGTGCAGGACTCGAAATGTCTCAATCTATCCGGCCACATAACGGTCAACCGT	600
Db	541	CATCAGACGGTGCAGGACTCGAAATGTCTCAATCTATCCGGCCACATAACGGTCAACCGT	600
Qy	601	ATGCTTTGGGATATGATGATGAACCTGCTGCGCTCAACACGGCCCTGTGTGTATCGACGTG	660
Db	601	ATGCTTTGGGATATGATGATGAACCTGCTGCGCTCAACACGGCCCTGTGTGTATCGACGTG	660
Qy	661	CTCCGGATCCACACAGCTGTCTGTGACATGTGTGGCGGGGCCCATTTGGGGATCTCGCGG	720
Db	661	CTCCGGATCCACACAGCTGTCTGTGACATGTGTGGCGGGGCCCATTTGGGGATCTCGCGG	720
Qy	721	GGTCTCGCTACTATTTCATGGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGCTACTTC	780
Db	721	GGTCTCGCTACTATTTCATGGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGCTACTTC	780
Qy	781	TTTGCTCCCTAATAG	795
Db	781	TTTGCTCCCTAATAG	795

RESULT 2  
AAL48914

AAI48914	standard; DNA; 795 BP.
AAI48914;	
24-OCT-2002	(first entry)
Hepatitis C virus clone HCC110A	E1 protein coding sequence.
Hepatitis C virus; HCV; E1 protein;	E2 protein; infection; gene;
virucide; immunostimulant; vaccine;	ds.
Hepatitis C virus.	
WO200255548-A2.	
18-JUL-2002.	
11-JAN-2002;	2002WO-EP000219.
11-JAN-2001;	2001US-0260669P.
30-AUG-2001;	2001US-0315768P.
(INNO-)	INNOGENETICS NV.

XX  
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 U; 0 Other;  
Query Match 100.0%; Score 795; DB 6; Length 795;  
Best Local Similarity 100.0%; Pred. No. 4.1e-204;  
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGGCGACCTCGTGGGGGTACATT	60
DB	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGGCGACCTCGTGGGGGTACATT	60
QY	61	CGGCTGCTGGCGGCCCGCCCTTAGGGGGCGCTGCCAAGGGCCCTGGCGCATGGGCGTCGGGGTT	120
DB	61	CGGCTGCTGGCGGCCCGCCCTTAGGGGGCGCTGCCAAGGGCCCTGGCGCATGGGCGTCGGGGTT	120
QY	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGGTGCTCTTTCTCTATCTTC	180
DB	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGGTGCTCTTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTTGCTGTCCTGTCTGACCGTTTCAGCTTCCGCTTATGAAGTGGCGAACGTG	240
DB	181	CTCTTGGCTTTTGCTGTCCTGTCTGACCGTTTCAGCTTCCGCTTATGAAGTGGCGAACGTG	240
QY	241	TCCGGGATGTACCATGTGTCAGAAACGACTGCTCCAACTCAAGCATTTGCTATGAGGCGACGG	300



ID AAL48939 standard; DNA; 2082 BP.  
 AC AAL48939;  
 DT 24-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.  
 XX  
 KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;  
 XX virucide; immunostimulant; vaccine; ds.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO20025548-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 11-JAN-2002; 2002WO-EP000219.  
 XX  
 PR 11-JAN-2001; 2001US-0260669P.  
 XX  
 PR 30-AUG-2001; 2001US-0315768P.  
 XX  
 XX (INNO-) INNOGENETICS NV.  
 PA  
 XX Maertens G, Bosman F, Buyse M;  
 PI P-PSDB; AAO18678.  
 XX  
 DR WPI; 2002-599657/64.  
 XX  
 DR P-PSDB; AAO18678.  
 XX  
 PT New therapeutic vaccine compositions comprising at least one purified  
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric  
 PT recombinant envelope protein E1 or E2, useful for immunizing humans from  
 PT HCV infection.  
 XX  
 PS Example 2; Page 206-209; 243pp; English.  
 XX  
 CC The present invention relates to new therapeutic vaccine compositions for  
 CC inducing Hepatitis C virus (HCV)-specific antibodies, comprising a  
 CC composition containing at least one purified recombinant HCV single or  
 CC specific oligomeric recombinant envelope proteins selected from an E1 and  
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are  
 CC useful for inducing HCV-specific antibodies or for immunising humans  
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as  
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody  
 CC tests, for raising antibodies, in the preparation of medicament, and for  
 CC in vitro monitoring of HCV disease or prognosis the response to  
 CC treatment of patients suffering from HCV infection. The present sequence  
 CC is a coding sequence described in the exemplification of the invention  
 XX  
 SQ Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 U; 0 Other;  
 Query Match 98.0%; Score 778.8; DB 6; Length 2082;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-199;  
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 4 TTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 63  
 4 TTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 63  
 64 CTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTTCG 123  
 64 CTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTTCG 123  
 124 GAGGACGGGTGAATATGATGCAACAGGAAATTTGCCGGTTCCTTCTATCTTCCTC 183  
 124 GAGGACGGGTGAATATGATGCAACAGGAAATTTGCCGGTTCCTTCTATCTTCCTC 183  
 184 TTGGCTTTCTGCTGCTGCTGATCCGCTTCAGCTTCGCTTATGAGTGGCAACGTGCC 243  
 184 TTGGCTTTCTGCTGCTGCTGATCCGCTTCAGCTTCGCTTATGAGTGGCAACGTGCC 243  
 244 GGGATGTATCATGTCAAGCAAGCATGCTCCAACTCAAGCAATTTGTATGAGGACGGGAC 303

Db 244 GGGATGTATCATGTCAAGCAAGCATGCTCCAACTCAAGCAATTTGTATGAGGACGGGAC 303  
 QY 304 ATGATCATGTCAACACCCCGGGTGGCTCGCTGGTTCGGGAGAACAACTCTTCCCGTGC 363  
 Db 304 ATGATCATGTCAACACCCCGGGTGGCTCGCTGGTTCGGGAGAACAACTCTTCCCGTGC 363  
 QY 364 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 423  
 Db 364 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 423  
 QY 424 CGAGCCACGTGCAATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTAAGTGGGG 483  
 Db 424 CGAGCCACGTGCAATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTAAGTGGGG 483  
 QY 484 GACCTCTGCGGATCTGCTTTCTCGCTCTCCAGCTGTTCACCATCTCGCTCCCGGCGAT 543  
 Db 484 GACCTCTGCGGATCTGCTTTCTCGCTCTCCAGCTGTTCACCATCTCGCTCCCGGCGAT 543  
 QY 544 GAGACGTGTCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603  
 Db 544 GAGACGTGTCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603  
 QY 604 GCTTGGGATATGATGATGAATCTGCTGCTCAACAGGCGCTTGGTGTATGCGAGCTGCTC 663  
 Db 604 GCTTGGGATATGATGATGAATCTGCTGCTCAACAGGCGCTTGGTGTATGCGAGCTGCTC 663  
 QY 664 CGGATCCCAACAGCTGTCTGCGACATGTGTGGGGGGGGCCCATTTGGGGAGTCTCTGGGGGT 723  
 Db 664 CGGATCCCAACAGCTGTCTGCGACATGTGTGGGGGGGGCCCATTTGGGGAGTCTCTGGGGGT 723  
 QY 724 CTCGCTACTATTTCCATGATGTCGGGAACTCGGGCTTAAGGTTTGTATGATGCTACTCTTT 783  
 Db 724 CTCGCTACTATTTCCATGATGTCGGGAACTCGGGCTTAAGGTTTGTATGATGCTACTCTTT 783  
 QY 784 GC 785  
 Db 784 GC 785  
 RESULT 5  
 ADD55555 standard; DNA; 2082 BP.  
 AC ADD55555;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Hepatitis C virus E1/E2 protein coding sequence #13.  
 KW Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;  
 KW liver fibrosis; ds; gene.  
 OS Hepatitis C virus.  
 XX  
 PN WO2003051912-A2.  
 XX  
 PD 26-JUN-2003.  
 XX  
 PF 18-DEC-2002; 2002WO-EP014480.  
 XX  
 PR 18-DEC-2001; 2001US-00020510.  
 XX  
 PR 16-OCT-2002; 2002US-0418358P.  
 XX  
 XX (INNO-) INNOGENETICS NV.  
 XX  
 XX Maertens G, Depla E, Bosman F;  
 PI P-PSDB; ADD55556.  
 XX  
 DR WPI; 2003-541632/S1.  
 XX  
 DR P-PSDB; ADD55556.  
 XX  
 PT New hepatitis C virus (HCV) vaccine composition, useful for reducing  
 PT liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.  
 XX  
 XX

PS Example 2; SEQ ID NO 47; 271bp; English.

CC The invention comprises an Hepatitis C virus (HCV) vaccine for reducing liver disease. The vaccine of the invention comprises an HCV E1 or E2 protein as an antigen. The HCV vaccine is useful for reducing liver disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The CC present DNA sequence encodes an HCV E1/E2 protein.

XX

Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 U; 0 Other;

Query Match 98.0%; Score 778.8; DB 9; Length 2082;

Best Local Similarity 99.7%; Pred. No. 1.3e-199;

Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 63

DB 4 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 63

QY 64 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123

DB 64 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123

QY 124 GAGGACGGCGTGAACATGATCAACAGGGAATTTGGCGGTTGCTTTCTCTATCTTCCTC 183

DB 124 GAGGACGGCGTGAACATGATCAACAGGGAATTTGGCGGTTGCTTTCTCTATCTTCCTC 183

QY 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

DB 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

QY 244 GGGATGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303

DB 244 GGGATGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303

QY 304 ATGATCATATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363

DB 304 ATGATCATATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363

QY 364 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 423

DB 364 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 423

QY 424 CGAGGCGACGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483

DB 424 CGAGGCGACGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483

QY 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543

DB 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543

QY 544 GAGACGCTGACGAGTTCGATATGCTCAATCTATCCCGGCGCACATPAAAGGCTACCGGTATG 603

DB 544 GAGACGCTGACGAGTTCGATATGCTCAATCTATCCCGGCGCACATPAAAGGCTACCGGTATG 603

QY 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663

DB 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663

QY 664 CGGATCCCAAGCTGCTGCGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723

DB 664 CGGATCCCAAGCTGCTGCGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723

QY 724 CTGCGCTACTATTCATGCTGGGGAACCTGGGCTAAAGGTTTGGATGATGATGATGATGATGATG 783

DB 724 CTGCGCTACTATTCATGCTGGGGAACCTGGGCTAAAGGTTTGGATGATGATGATGATGATGATG 783

QY 784 GC 785

DB 784 GC 785

ID XX AAT12973 standard; DNA; 2086 BP.

AC AAT12973;

DT 24-SEP-1996 (first entry)

XX HCV E1 construct HCCI65.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.

OS Hepatitis C virus.

FN MO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95MO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman P, De Martynoff G, Buyse M;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of disulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques

XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 U; 0 Other;

Query Match 98.0%; Score 778.8; DB 2; Length 2086;

Best Local Similarity 99.7%; Pred. No. 1.3e-199;

Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 63

DB 4 TTGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 63

QY 64 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123

DB 64 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123

QY 124 GAGGACGGCGTGAACATGATCAACAGGGAATTTGGCGGTTGCTTTCTCTATCTTCCTC 183

DB 124 GAGGACGGCGTGAACATGATCAACAGGGAATTTGGCGGTTGCTTTCTCTATCTTCCTC 183

QY 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

DB 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

244 GGGATGTACCATGTACGAGCACTGCTCTCAACTCAAGCATTTGTGTATGAGCAGCGGAC 303  
 244 GGGATGTACCATGTACGAGCACTGCTCTCAACTCAAGCATTTGTGTATGAGCAGCGGAC 303  
 304 ATGATCATGCACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
 304 ATGATCATGCACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
 364 TGGGTAGGGCTCACCCCGGCTGGCAGCTAGGAAAGCCAGCGTCCCAACACGACCAATA 423  
 364 TGGGTAGGGCTCACCCCGGCTGGCAGCTAGGAAAGCCAGCGTCCCAACACGACCAATA 423  
 424 CGAGCCACGCTGATTTGCTGCTGGGGGGGCTGCTTTCTGTTCCGCTATGTAACGTGGG 483  
 424 CGAGCCACGCTGATTTGCTGCTGGGGGGGCTGCTTTCTGTTCCGCTATGTAACGTGGG 483  
 484 GACCTCTGGGATCTGCTCTGCTCCAGCTGTTCCAGCTGTTCCAGCTGTTCCAGCTGTT 543  
 484 GACCTCTGGGATCTGCTCTGCTCCAGCTGTTCCAGCTGTTCCAGCTGTTCCAGCTGTT 543  
 544 GAGAGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603  
 544 GAGAGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603  
 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
 664 CGATATCCCAAGCTGTGTGAGCATGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 723  
 664 CGATATCCCAAGCTGTGTGAGCATGTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 723  
 724 CTGCGCTACTATTCCATGTTGGGAACTGGGGCTAAGGTTTTCATTTGTATGATGATCTTT 783  
 724 CTGCGCTACTATTCCATGTTGGGAACTGGGGCTAAGGTTTTCATTTGTATGATGATCTTT 783  
 784 GC 785  
 784 GC 785

RESULT 7  
 AAT12974  
 ID AAT12974 standard; DNA; 2433 BP.  
 XX AC AAT12974;  
 XX DT 25-SEP-1996 (first entry)  
 XX DE HCV E1 construct HCCI66.  
 XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;  
 XX KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;  
 XX KW ss.  
 XX OS Hepatitis C virus.  
 XX PN W09604385-A2.  
 XX PD 15-FEB-1996.  
 XX PF 31-JUL-1995; 95WO-EP003031.  
 XX PR 29-JUL-1994; 94EP-00870132.  
 XX PA (INNO-) INNOGENETICS NV.  
 XX PI Maertens G, Bosman F, De Martynoff G, Buysse M;  
 XX WPI; 1996-129401/13.  
 PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins

PT - in presence of di: sulphide bond cleavage agent, to produce proteins  
 PT suitable for direct use in vaccines or diagnostic assays of HCV.  
 XX Claim 23; Fig 21; 146pp; English.  
 XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1  
 CC and E2 protein coding sequence constructs. These sequences are included  
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.  
 CC The recombinant proteins can then be isolated using a method of the  
 CC invention. In the method, the envelope proteins are purified by carrying  
 CC out a disulphide bond cleavage, or a reduction step with a disulphide  
 CC bond cleavage agent, after lysis of recombinant host cells. The  
 CC constructs containing the purified HCV envelope proteins can be used for  
 CC vaccinating humans against HCV, for in vitro detection of HCV antibodies  
 CC in a sample, and in a serotyping assay for detecting one or more  
 CC serological types of HCV present in a biological sample. The constructs  
 CC can also be immobilised on a solid substrate and incorporated into a  
 CC reversed phase hybridisation assay for determining the presence or the  
 CC genotype of HCV. The new purification method preserves the conformation  
 CC of the recombinantly expressed E1, E2 and E1/E2, and eliminates  
 CC contaminating proteins. Antigens isolated using this method are more  
 CC reactive with human sera than those isolated by known techniques  
 XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 U; 0 Other;  
 SQ

Query Match 98.0%; Score 778.8; DB 2; Length 2433;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-199;  
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAGGTGATCATGATACCTTATACATCGCGGCTTGGCGAGCTCGTGGGGTACATTCG 63  
 DB 355 TTGGGTAGGTGATCATGATACCTTATACATCGCGGCTTGGCGAGCTCGTGGGGTACATTCG 414  
 QY 64 CTCGTGCGCGCGCCCTAGGGGGCGCTGCGAGGCGCTGCGAGTGGCGGTTCG 123  
 DB 415 CTCGTGCGCGCGCCCTAGGGGGCGCTGCGAGGCGCTGCGAGTGGCGGTTCG 474  
 QY 124 GAGGACGGCGTGAATCATGATGCAACAGGGAATTTGCCGGTGTCTTTCTATTTCTCTC 183  
 DB 475 GAGGACGGCGTGAATCATGATGCAACAGGGAATTTGCCGGTGTCTTTCTATTTCTCTC 534  
 QY 184 TTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243  
 DB 535 TTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594  
 QY 244 GGGATGTACCATGTACGAGCACTGCTCTCAACTCAAGCATTTGTGTATGAGCAGCGGAC 303  
 DB 595 GGGATGTACCATGTACGAGCACTGCTCTCAACTCAAGCATTTGTGTATGAGCAGCGGAC 654  
 QY 304 ATGATCATGCACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
 DB 655 ATGATCATGCACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 714  
 QY 364 TGGGTAGGGCTCACCCCGGCTGGCAGCTAGGAAAGCCAGCGTCCCAACACGACCAATA 423  
 DB 715 TGGGTAGGGCTCACCCCGGCTGGCAGCTAGGAAAGCCAGCGTCCCAACACGACCAATA 774  
 QY 424 CGAGCCACGCTGATTTGCTGCTGGGGGGGCTGCTTTCTGTTCCGCTATGTAACGTGGG 483  
 DB 775 CGAGCCACGCTGATTTGCTGCTGGGGGGGCTGCTTTCTGTTCCGCTATGTAACGTGGG 834  
 QY 484 GACCTCTGGGATCTGCTCTGCTCCAGCTGTTCCAGCTGTTCCAGCTGTTCCAGCTGTT 543  
 DB 835 GACCTCTGGGATCTGCTCTGCTCCAGCTGTTCCAGCTGTTCCAGCTGTTCCAGCTGTT 894  
 QY 544 GAGAGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603  
 DB 895 GAGAGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 954  
 QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663  
 DB 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014



DR WPI; 2002-599657/64.  
 XX P-PSDB; AAO18679.  
 XX New therapeutic vaccine compositions comprising at least one purified  
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric  
 FT recombinant envelope protein E1 or E2, useful for immunizing from  
 PT HCV infection.  
 XX  
 XX Example 2; Page 212-215; 243pp; English.  
 XX  
 CC The present invention relates to new therapeutic vaccine compositions for  
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a  
 CC composition containing at least one purified recombinant HCV single or  
 CC specific oligomeric recombinant envelope proteins selected from an E1 and  
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are  
 CC useful for inducing HCV-specific antibodies or for immunising humans  
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as  
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody  
 CC tests, for raising antibodies, in the preparation of medicament, and for  
 CC in vitro monitoring of HCV disease or for prophylaxis of the response to  
 CC treatment of patients suffering from HCV infection. The present sequence  
 CC is a coding sequence described in the exemplification of the invention  
 XX  
 XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 U; 0 Other;  
 SQ  
 Query Match 96.6%; Score 767.8; DB 6; Length 2434;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-196;  
 Matches 780; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 4 TTGGTAAAGTCAATGATACCTTATACATCGCGCTTGGCGACCTCGTGGGGTACATTCG 63  
 Db 355 TTGGTAAAGTCAATGATACCTTATACATCGCGCTTGGCGACCTCGTGGGGTACATTCG 414  
 Qy 64 CTGTCGCGGCGCCCTAGGCGGCGCTGCGAGGCGCTGCGGCGAGCGTCCGGGTTG 123  
 Db 415 CTGTCGCGGCGCCCTAGGCGGCGCTGCGAGGCGCTGCGGCGAGCGTCCGGGTTG 474  
 Qy 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 183  
 Db 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 534  
 Qy 184 TTGGTCTTGTCTCTG-TCTGACCGTTCCAGCTTCGCGCTTATGAAGTCCGCAAGTGT 242  
 Db 535 TTGGTCTTGTCTCTG-TCTGACCGTTCCAGCTTCGCGCTTATGAAGTCCGCAAGTGT 594  
 Qy 243 CGGATGTACATGTCTGACGACGCTGCTCAACTCAAGCATTTGTATGAGCGACGGA 302  
 Db 595 CGGATGTACATGTCTGACGACGCTGCTCAACTCAAGCATTTGTATGAGCGACGGA 654  
 Qy 303 CATGATCATGACACCCCGGTTGCGTCCGTTTCCGAGAACAACTCTTCCCGCTG 362  
 Db 655 CATGATCATGACACCCCGGTTGCGTCCGTTTCCGAGAACAACTCTTCCCGCTG 714  
 Qy 363 CTGGTAGCGCTCACCCCGACCTGCGAGTACGAGTACGAGTCCGCCACCAACGACAT 422  
 Db 715 CTGGTAGCGCTCACCCCGACCTGCGAGTACGAGTACGAGTCCGCCACCAACGACAT 774  
 Qy 423 ACAGCGCCACGTCGATTTGCTGTTGGGGGCGCTTCTCTGCGCTATGACGTGG 482  
 Db 775 ACAGCGCCACGTCGATTTGCTGTTGGGGGCGCTTCTCTGCGCTATGACGTGG 834  
 Qy 483 GGACCTCTGCGGATCTGCTCTCTGTCCTCCAGCTGTTTCCACCATCTCGCTCCCGGCA 542  
 Db 835 GGACCTCTGCGGATCTGCTCTCTGTCCTCCAGCTGTTTCCACCATCTCGCTCCCGGCA 894  
 Qy 543 TGAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGGTCCCGTAT 602  
 Db 895 TGAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGGTCCCGTAT 954  
 Qy 603 GCGTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662  
 Db 955 GCGTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014

Qy 663 CCGGATCCCAACAGCTGCTGTCGACATGCTGGCGGGGCCCATTTGGGAGTCTCTGCGGG 722  
 Db 1015 CCGGATCCCAACAGCTGCTGTCGACATGCTGGCGGGGCCCATTTGGGAGTCTCTGCGGG 1074  
 Qy 723 TCTCGCTTACTATTCATGCTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTT 782  
 Db 1075 CCTCGCTTACTATTCATGCTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTT 1134  
 Qy 783 TGC 785  
 Db 1135 TGC 1137  
 RESULT 10  
 ABK91431  
 ID ABK91431 standard; DNA; 9605 BP.  
 AC ABK91431;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate DNA mutant 8.  
 XX  
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
 OS Hepatitis C virus.  
 XX  
 XX Synthetic.  
 XX  
 Key Location/Qualifiers  
 CDS 342..9374  
 FT /\*tag= a  
 FT /product= "HCV polyprotein"  
 FT /note= "The polyprotein consists of the Core, E1, E2, P7,  
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT mutation replace(6934,T)  
 FT /\*tag= b  
 FT  
 FT WO200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 PP 16-JAN-2002; 2002MO-BP000526.  
 XX  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) IST RICERCHI BIOL MOLECOLARE ANGELETTI.  
 XX  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 WPI; 2002-599793/64.  
 XX  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 XX  
 XX Claim 9; Page; 69pp; English.  
 PS  
 XX  
 CC The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing an HCV (hepatitis C virus)  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV



CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
XX appearing as ABK91411 and the information in Claim 9

SQ Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 U; 0 Other;

Query Match 88.5%; Score 703.6; DB 6; Length 9605;  
Best Local Similarity 93.7%; Pred. No. 4e-179;  
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAGCTCATGATACCTTACATGCGCTTGCCGACCTCGTGGGTACATCCG 63  
DB |||  
696 TTGGGTAAGCTCATGATACCTTACATGCGCTTGCCGACCTCGTGGGTACATCCG 755  
QY 64 CTGCTGCGCGCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTTCG 123  
DB |||  
756 CTGCTGCGCGCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTTCG 815  
QY 124 GAGACGCGGTGAATGATGACAGGAAATTTGGCGGTTCCTCTTCTATCTCTC 183  
DB |||  
816 GAGACGCGGTGAATGATGACAGGAAATTTGGCGGTTCCTCTTCTATCTCTC 875  
QY 184 TTGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTATGAGTGCGCAAGTGC 243  
DB |||  
876 TTGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTATGAGTGCGCAAGTGC 935  
QY 244 GGGATGTACCATGTTCAGAACGATGCTTCAAGCATTTGTATGAGGACGCGGAC 303  
DB |||  
936 GGGATGTACCATGTTCAGAACGATGCTTCAAGCATTTGTATGAGGACGCGGAC 995  
QY 304 ATGATCATGACACCCCGGGTGGTGGTCTTGGCGTTCGGGAGAACATCTTCCGCTGC 363  
DB |||  
996 ATGATCATGACACCCCGGGTGGTGGTCTTGGCGTTCGGGAGAACATCTTCCGCTGC 1055  
QY 364 TGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCACACACACATA 423  
DB |||  
1056 TGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCACACACATA 1115  
QY 424 CGACGCGCACGCTCGATTGCTCGTGGGCGGCTGCTTCTGTTCCGCTATGATGAGTGGG 483  
DB |||  
1116 CGACGCGCACGCTCGATTGCTCGTGGGCGGCTGCTTCTGTTCCGCTATGATGAGTGGG 1175  
QY 484 GACCTTGGCGATCTGCTTCTGCTGCTCCAGCTGTTCCACATCTCCCTTCGCGCGCAT 543  
DB |||  
1176 GATCTCTGCGGATCTGTTTCTGCTGCTCCAGCTGTTCCACATCTCCCTTCGCGCGCAT 1235  
QY 544 GAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCGCACATACGGTCAACGCTATG 603  
DB |||  
1236 GAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCGCACATACGGTCAACGCTATG 1295  
QY 604 GCTTGGGATATGATGAACTGCTGCTTCAACGCGCTGCTGCTGCTGCTGCTGCTGCT 663  
DB |||  
1296 GCTTGGGATATGATGAACTGCTGCTTCAACGCGCTGCTGCTGCTGCTGCTGCTGCT 1355  
QY 664 CGGATCCACAGCTGCTGCGACATGCTGGCGGGGCGCCATTTGGGAGTCTTGGCGGGT 723  
DB |||  
1356 CGGATCCACAGCTGCTGCGACATGCTGGCGGGGCGCCATTTGGGAGTCTTGGCGGGT 1415  
QY 724 CTGCTTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGATGATCTCTTT 783  
DB |||  
1416 CTGCTTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGATGATCTCTTT 1475  
QY 784 GC 785  
DB |||  
1476 GC 1477

RESULT 11  
ABK91424  
ID ABK91424 standard; DNA; 9605 BP.  
XX  
AC ABK91424;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus Con 1 isolate DNA mutant 1.  
XX  
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
XX  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
CDS 342..9374  
/tag= a  
/product= "HCV polyprotein"  
/note= "The polyprotein consists of the Core, E1, E2, P7,  
NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
FT mutation  
replace(3625,G)  
/tag= b  
XX  
PN WO200259321-A2.  
XX  
PD 01-AUG-2002.  
XX  
PP 16-JAN-2002; 2002WO-EP000526.  
XX  
PR 23-JAN-2001; 2001US-0263479P.  
XX  
PA (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.  
XX  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
DR WPI; 2002-599793/64.  
XX  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 9; Page; 69pp; English.  
XX  
CC The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replicon RNA and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABK91411 and the information in Claim 9  
XX  
SQ Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 88.5%; Score 703.6; DB 6; Length 9605;

Best Local Similarity 93.7%; Pred. No. 4e-179;  
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 63  
DB TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 755

QY 64 CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 123  
DB CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 815

QY 124 GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 183  
DB GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 875

QY 184 TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 243  
DB TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 876

QY 244 GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 303  
DB GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 936

QY 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
DB ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996

QY 364 TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423  
DB TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056

QY 424 CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
DB CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116

QY 484 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543  
DB GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176

QY 544 GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 603  
DB GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 1236

QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
DB GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1296

QY 664 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723  
DB CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356

QY 724 CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 783  
DB CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 1416

QY 784 GC 785  
DB GC 1477

RESULT 12  
ID ABK91429  
XX ABK91429 standard; DNA; 9605 BP.  
XX ABK91429;  
XX AC  
XX DB  
XX DT  
XX DE  
XX KW  
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

15-NOV-2002 (first entry)  
Hepatitis C virus Con 1 isolate DNA mutant 6.

QY 4 TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 63  
DB TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 755

QY 64 CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 123  
DB CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 815

QY 124 GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 183  
DB GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 875

QY 184 TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 243  
DB TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 876

QY 244 GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 303  
DB GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 936

QY 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
DB ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996

QY 364 TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423  
DB TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056

QY 424 CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
DB CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116

QY 484 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543  
DB GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176

QY 544 GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 603  
DB GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 1236

QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
DB GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1296

QY 664 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723  
DB CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356

QY 724 CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 783  
DB CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 1416

QY 784 GC 785  
DB GC 1477

Query Match 88.5%; Score 703.6; DB 6; Length 9605;  
Best Local Similarity 93.7%; Pred. No. 4e-179;  
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 63  
DB TTGGGTAAGGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCGG 755

QY 64 CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 123  
DB CTGCTCGGCGCCCTTACGAGGCGCTTCGAGGCGCTTCGCGACCTTCGCGGCTTCG 815

QY 124 GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 183  
DB GAGGACGGCGTGAATATGCAACAGAGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 875

QY 184 TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 243  
DB TTGGCTTTGCTGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTCGAGGCTTC 876

QY 244 GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 303  
DB GGGATGTACCATGTCAAGCAAGCACTCTCCAACTCAAGCAATTTGTATGAGGCGAGCGGAC 936

QY 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
DB ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996

QY 364 TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423  
DB TGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056

QY 424 CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
DB CGAGCGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116

QY 484 GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543  
DB GACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176

QY 544 GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 603  
DB GAGACGGTGCAGACCTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGGTACCGGTATG 1236

QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663  
DB GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1296

QY 664 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723  
DB CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356

QY 724 CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 783  
DB CTGCGCTACTATTCCATGCTGGGGAAGCTGGGCTAAAGCTTTTGATGATGATGATGATGATG 1416

QY 784 GC 785  
DB GC 1477

Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;

Query Match 88.5%; Score 703.6; DB 6; Length 9605;  
Best Local Similar

Db 816 GAGGACGGGTGAACATATCAACAGGGAATCTGCCGGTGTCTCTTCTATCTTCCTT 875  
 QY 184 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 243  
 Db 876 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 935  
 QY 244 GGGATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303  
 Db 936 GGAAGTGTACCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 995  
 QY 304 ATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 363  
 Db 996 ATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 1055  
 QY 364 TTGGTGTAGCTTCACTCCAGCTGCTGAGTGTAGGAGCGGAGCGGAGCGGAGCGGAG 423  
 Db 1056 TTGGTGTAGCTTCACTCCAGCTGCTGAGTGTAGGAGCGGAGCGGAGCGGAGCGGAG 1115  
 QY 424 CGACGCGACGCTGATTTGCTCTGTTGGGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGG 483  
 Db 1116 CGACGCGACGCTGATTTGCTCTGTTGGGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGG 1175  
 QY 484 GACCTCTGCGAGTGTCTCTGCTCTCCAGCTGTTTACCATCTGCTGCTGCGGGGAT 543  
 Db 1176 GATCTCTGCGAGTGTCTCTGCTCTCCAGCTGTTTACCATCTGCTGCTGCGGGGAT 1235  
 QY 544 GAGAGCTGTGAGGACTGCTTCAATCTATCTCCGCGGCGACATACAGGCTACCGGTATG 603  
 Db 1236 GAGAGCTGTGAGGACTGCTTCAATCTATCTCCGCGGCGACATACAGGCTACCGGTATG 1295  
 QY 604 GCTTGGGATATGATGAACTGCTGCTTCAATCTATCTCCGCGGCGACATACAGGCTACCG 663  
 Db 1296 GCTTGGGATATGATGAACTGCTGCTTCAATCTATCTCCGCGGCGACATACAGGCTACCG 1355  
 QY 664 CGATCCCAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 723  
 Db 1356 CGATCCCAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1415  
 QY 724 CTCGCTTACTTATCCATGTTGGGAACTGCGGCTAAGGTTTGTATGATGCTACTCTTT 783  
 Db 1416 CTCGCTTACTTATCCATGTTGGGAACTGCGGCTAAGGTTTGTATGATGCTACTCTTT 1475  
 QY 784 GC 785  
 Db 1476 GC 1477

RESULT 13  
 ID ABK91432  
 AC ABK91432; standard; DNA; 9605 BP.  
 DT 15-NOV-2002 (first entry)  
 XX Hepatitis C virus Con 1 isolate DNA mutant 9.  
 XX Hepatitis C virus.  
 XX Synthetic.  
 XX Key  
 XX CDS  
 XX Location/Qualifiers  
 342..9374  
 /tag= a  
 /product= "HCV polyprotein"  
 /note= "The polyprotein consists of the Core, E1, E2, P7,  
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 mutation replace(6936,G)  
 /tag= b

PN WO200259321-A2.  
 XX 01-AUG-2002.  
 XX 16-JAN-2002; 2002WO-EP000526.  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 XX WPI; 2002-599793/64.  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 XX Claim 9; Page; 69pp; English.  
 XX The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABK91411 and the information in Claim 9  
 XX SQ Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 88.5%; Score 703.6; DB 6; Length 9605;  
 Best Local Similarity 93.7%; Pred. No. 4e-179;  
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
 QY 4 TTGGTAAAGTTCATGATACCTTACATGCGGCTTGGCGGCTGCGGGGTACATTCCG 63  
 Db 696 TTGGTAAAGTTCATGATACCTTACATGCGGCTTGGCGGCTGCGGGGTACATTCCG 755  
 QY 64 CTCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 123  
 Db 756 CTCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815  
 QY 124 GAGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183  
 Db 816 GAGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875  
 QY 184 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 243  
 Db 876 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 935  
 QY 244 GGGATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303  
 Db 936 GGAAGTGTACCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 995  
 QY 304 ATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 363  
 Db 996 ATGATCATGTCACGACGACTGCTTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 1055

QY 364 TGGGTAGCGCTCACCCACGCTCGCAGCTAGGAAAGCCAGAGCTCCGCCACCAACGACAATA 423  
 Db 1056 TGGGTAGCGCTCACTCCCAAGCTCGCGGCCAGGACGCTAGCGTCCCACTACGACGATA 1115  
 QY 424 CGAGCCAGCTCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGG 483  
 Db 1116 CGAGCCAGCTCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGA 1175  
 QY 484 GACCTCTGGGATCTGCTTCTGCTCTCCAGCTGTCACCATCTCGCTCGCGGCAT 543  
 Db 1176 GATCTCTGGGATCTGCTTCTGCTCTCCAGCTGTCACCATCTCGCTCGCGGCAC 1235  
 QY 544 GAGACGCTGAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCACCGTATG 603  
 Db 1236 GAGACAGTACAGGACTGCAATTGCTCAATATATCCCGGCCAGTACAGGTACCGTATG 1295  
 QY 604 GCTTGGGATATGATGATGATGCTGCTGCTTACAGCGGCTGCTGCTGATGACGTGCTC 663  
 Db 1296 GCTTGGGATATGATGATGATGCTGCTGCTTACAGCGGCTGCTGCTGATGACGTGCTC 1355  
 QY 664 CGATCCCAACAGCTGCTGTCGACATGCTGGGGGGCCCATTTGGGAGTCTTGGGGGT 723  
 Db 1356 CGATCCCAACAGCTGCTGTCGACATGCTGGGGGGCCCATTTGGGAGTCTTGGGGGT 1415  
 QY 724 CTCGCTACTATTCCATGCTGGGGAAGTGGGCTTAAGGTTTGTGATGCTGCTACTCTTT 783  
 Db 1416 CTGCTACTATTCCATGCTGGGGAAGTGGGCTTAAGGTTTGTGATGCTGCTACTCTTT 1475  
 QY 784 GC 785  
 Db 1476 GC 1477  
 RESULT 14  
 ID ABK91411  
 AC ABK91411;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate DNA.  
 XX  
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Location/Qualifiers  
 FT CDS 342..9374  
 FT /\*tag= a  
 FT /product= "HCV polyprotein"  
 FT /note= "The polyprotein consists of the Core, E1, E2, E7,  
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 XX  
 FN WO200259321-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 16-JAN-2002; 2002WO-EP000526.  
 XX  
 PR 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
 PA De Francesco R, Migliaccio G, Paonessa G;  
 XX  
 XX WPI; 2002-599793/64.  
 DR P-PSDB; ABG32451.  
 XX  
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV

PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 XX expression.  
 PS Claim 9; Page 36-39; 69pp; English.  
 XX  
 CC The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1,  
 CC used as a basis for the adaptive mutations of the invention  
 XX  
 SQ Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;  
 Query Match 88.5%; Score 703.6; DB 6; Length 9605;  
 Best Local Similarity 93.7%; Pred. No. 4e-179;  
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
 QY 4 TTGGGTAAAGTTCATCGATACCTTACATGGGCTTCGCGGACCTTCGTGGGTACATTCOG 63  
 Db 696 TTGGGTAAAGTTCATCGATACCTTACATGGGCTTCGCGGATCTCATGGGTACATTCOG 755  
 QY 64 CTCGTGCGCGCCCTTAGCGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCGTTCTG 123  
 Db 756 CTCGTGCGCGCCCTTAGCGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCGTTCTG 815  
 QY 124 GAGGACGGGTGAATATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCCTC 183  
 Db 816 GAGGACGGGTGAATATATGCAACAGGGAATCTGCCGGTTCCTCTTCTATCTTCCTC 875  
 QY 184 TTGGGTTCCTGCTCTGCTGACCGCTTCAGCTTCGCTTATGAAAGTGCACAGTGTCC 243  
 Db 876 TTGGGTTCCTGCTCTGCTGACCGCTTCAGCTTCGCTTATGAAAGTGCACAGTGTCC 935  
 QY 244 GGGATGTACCATGTCAAGAACGACCTGCTCAAGCAATGTGTATGAGCGACGGAC 303  
 Db 936 GGAAGTGTACCATGTCAAGAACGACCTGCTCAAGCAAGCAATGTGTATGAGCGACGGAC 995  
 QY 304 ATGATCATGACACACCCCGGGTGCCTGCGCTCGGTTCGGGAGAACAACTCTTCCTCGC 363  
 Db 996 ATGATCATGACACACCCCGGGTGCCTGCGCTCGGTTCGGGAGAACAACTCTTCCTCGC 1055  
 QY 364 TGGGTAGCGCTCACCCCAAGCTCGCAGTACGAAAGCGCGAGCGTTCGCCACCAACATA 423  
 Db 1056 TGGGTAGCGCTCACCTCCCAAGCTCGCGGCCAGAACGCTAGCGTTCGCCACCAATA 1115  
 QY 424 CGAGCCAGCTCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGG 483  
 Db 1116 CGAGCCAGCTCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGA 1175  
 QY 484 GACCTCTGGGATCTGCTTCTGCTCTCCAGCTGTCACCATCTCGCTCGCGGCAT 543  
 Db 1176 GATCTCTGGGATCTGCTTCTGCTCTCCAGCTGTCACCATCTCGCTCGCGGCAC 1235  
 QY 544 GAGACGCTGAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCACCGTATG 603  
 Db 1236 GAGACAGTACAGGACTGCAATTGCTCAATATATCCCGGCCAGTACAGGTACCGTATG 1295



Db        ||  
         1476 GC 1477

Search completed: June 16, 2004, 09:36:53  
Job time : 301.119 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 2137.22 Seconds  
(without alignments)  
11108.065 Million cell updates/sec

Title: US-09-899-303A-5

Perfect score: 795

Sequence: 1 ATGTTGGTAAAGTCATCGA.....TACTTTTGCTCCTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
c 1	115.6	14.5	488	9	AV755731
c 2	91.6	11.5	492	9	AV758366
c 3	47.2	5.9	1201	13	BX356664
4	43.4	5.5	925	29	CNS0091P

C	5	43	5.4	886	14	CK209050
	6	42.8	5.4	423	9	AA459034
C	7	41.6	5.2	1201	9	AL513886
	8	41.2	5.2	1195	28	BL0902
	9	40.6	5.1	402	9	AV392783
	10	40.6	5.1	551	9	AV392165
	11	40.6	5.1	552	12	BI996341
	12	40.6	5.1	584	12	BI727879
	13	40	5.0	525	10	BE337089
	14	40	5.0	534	14	CF846043
	15	40	5.0	671	12	BI723733
C	16	39.8	5.0	608	14	CB640103
C	17	39.8	5.0	738	14	CB668031
	18	39.8	5.0	969	12	BQ017656
C	19	39.6	5.0	675	12	BQ045314
	20	39.2	4.9	571	12	BM692316
	21	39.2	4.9	645	29	CNS01213
	22	39.2	4.9	1102	13	BQ069967
C	23	39	4.9	359	12	BQ252669
	24	39	4.9	375	12	BQ246716
	25	39	4.9	621	14	CA816001
	26	39	4.9	624	12	BI723734
	27	39	4.9	739	29	CG239174
C	28	39	4.9	744	29	CC605518
C	29	39	4.9	828	29	CG381985
	30	39	4.9	840	28	CC335916
	31	39	4.9	841	29	CC605508
	32	39	4.9	873	14	CD446071
C	33	39	4.9	1029	29	CC625582
	34	38.8	4.9	590	14	CA659369
	35	38.6	4.9	693	13	CA160850
	36	38.4	4.8	702	14	CD432549
	37	38.4	4.8	740	12	BJS36071
	38	38.4	4.8	970	29	CNS010C9
C	39	38.4	4.8	987	29	CNS015VX
C	40	38.2	4.8	1148	14	CK208676
C	41	38.2	4.8	1270	12	BG968359
C	42	38	4.8	354	14	CB966525
C	43	38	4.8	742	13	BQ752673
	44	37.8	4.8	435	13	C72860
	45	37.8	4.8	533	28	CC010084

#### ALIGNMENTS

RESULT 1	AV755731/c	488 bp	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	cdna clone BMFAR03 5', mRNA	sequence.
DEFINITION	AV755731	EST.		
ACCESSION	AV755731	GI:10913579		
VERSION	AV755731.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.			
TITLE	Homo sapiens cDNA BM clones			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.			



FEATURES  
source

Location/Qualifiers  
1..488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BMFAK03"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/clone\_lib="BM"  
/note="Vector: pTriplEx2; Site\_1: sfiIA; Site\_2: sfiIB"

## ORIGIN

Query Match 14.5%; Score 115.6; DB 9; Length 488;  
Best Local Similarity 64.5%; Pred. No. 1.6e-15;  
Matches 205; Conservative 0; Mismatches 109; Indels 4; Gaps 2;

Qy 445 GTTGGGGGGCTTTCTTCCTCGATGAGTGGGGGACCTCTGGGATCTGCTTC 504  
Db 472 GTGGTGCACACTCGCTCTGCTCAGCTCTCTACGTTGGGACCTCTGGGACGGAGTGATG 413

Qy 505 CTGCTCTCCAGCTGTTTCAACATCTCGCTCGCGGCATGAGACGCTGCAGGACTGCAAT 564  
Db 412 CTTCAGTTGAGCTG---ATCATCTGGCTCAGCACCATGATTTGTGCTGATGCAATGCAAC 356

Qy 565 TGCTCAATCTATCCGGGCACATAACGGGTCAACGGTATG-GCTTGGGATATGATGATGAA 623  
Db 355 TGCTCCATCTATCTCGCGCCATCACTGGACACCGGTATGAGCATGGGACATGATGAA 296

Qy 624 CTGTCGGCTACAAAGCGGCTGGTATCGGAGCTGCTCCGGATCCCAAGCTGTGTGT 683  
Db 295 CTGGTGTGTCACCGCTGCTGATGATCATGCGCGTACGCAATGCGCGTTCCTGAGGTATCAT 236

Qy 684 GGCATGTGTGGCGGGCGCCATGTCGGGAGTCTCGCGGCTCTCGCTACTATTTCATGTT 743  
Db 235 AGATATCATCAGCGGGCTCACTGGCGGTGATGTTGGCTTAGCTTCTCTATGCA 176

Qy 744 GGGGAACCTGGGCTAAGGT 761  
Db 175 GGGAGCGTGGCGCAAGT 158

## RESULT 2

AV758366/c 492 bp mRNA linear EST 19-OCT-2000  
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.

DEFINITION AV758366  
ACCESSION AV758366  
VERSION AV758366.1 GI:10916214

KEYWORDS  
SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 492)  
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,  
Gu,Y., Li,N., Qian,B., Liu,P., Qu,J., Gao,X., Cheng,Z., Xu,Z.,  
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and  
Chen,Z.

Homo sapiens cDNA BM clones

Unpublished (2000)

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

## FEATURES

## source

Location/Qualifiers  
1..492  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="BMFAK03"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/clone\_lib="BM"  
/note="Vector: pTriplEx2; Site\_1: sfiIA; Site\_2: sfiIB"

## ORIGIN

Query Match 11.5%; Score 91.6; DB 9; Length 492;  
Best Local Similarity 60.9%; Pred. No. 4.1e-10;  
Matches 184; Conservative 0; Mismatches 114; Indels 4; Gaps 2;

Qy 461 TCTGTTCCGCTATGTAAGTGGGGACCTCTCTGGGATCTGCTTCTTCTGCTTCCAGCTGT 520  
Db 457 TGTGATCAGCTCACTACGCTGTGGACCTCTGCGTTGGGTATCGCTTGCAGCCCACTG- 399

Qy 521 TCACCATCTCGCTCGCGGCATGAGACGCTGCAGACCTGCAATGCTCAATCTATCCCG 580  
Db 398 --ATTATCTCTCAGCAGCAACATTTGTTGCAAGATGCAACTGCTCATTTCTATCCCTG 341

Qy 581 GCCACATAACGGGTC-ACCGTATGGCTTGGGATATGATGATGAACCTGTCGCTCAACG 639  
Db 340 GCTGATCTACTGATCTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 281

Qy 640 GCCCTGTTGTTATCCGAGCTCTCCGATCCCAAGCTGTCGTGACATGTTGGCGGGG 699  
Db 280 TCCATGATACTGGCGTACGCAATGCGGTTCTCTGAAGTCTCTCATAGATATCATAGCTGG 221

Qy 700 GCCCATTTGGGAGTCTTGGCGGCTCTCGCTACTATTTCATGTTGGGAACTGGGCTAAG 759  
Db 220 GCACACTGCGGCTCATGTTGCGGCTCAGCTTACTTCACAATGAGGAGGCTTGGCCAA 161

Qy 760 GT 761  
Db 160 GT 159

## RESULT 3

## BX356664

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## CONTACT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT



## COMMENT

Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_estes@usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [65,694].  
Plate: L5B016 row: D column: 09.

## FEATURES

source

1. 886  
Location/Qualifiers  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum PGAS: Library 5 GATE 7"

/note="Vector: pCMV.SPORT6; Crown and developmental stages  
of spike formation in wheat cultivar Norstar. 4 mRNA  
populations were combined before constructing the library.  
The first mRNA population is from 1cm crown sections after  
30 days of cold acclimation. The second is from 1cm crown  
sections after 11 days of deacclimation (before  
deacclimation plants were fully vernalized for 49 days).  
The third is from different developmental stages of spike  
formation (5 to 50mm) that still have not emerged from the  
leaf (dissection required). The last is from different  
developmental stages of spike and seed formation after  
having emerged from the leaf (visible). First strand  
synthesis in this library was done in the presence of  
methylated dCTP thereby protecting from internal cleavage  
with NotI."

## ORIGIN

Query Match 5.4%; Score 43; DB 14; Length 886;  
Best Local Similarity 54.8%; Pred. No. 48;  
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 11 AGGTGATGATACCTTATGATGGCTTGGCGGACCTCGTGGGTACATTCGGCTGTCG 70  
Db 669 AGGCTTCGACGACACAGCGGGCGCGGAGACATCGCGGCGACGCGCTCCGCGACA 610

Qy 71 CGGCCCCCTAGGGGGCGCTGCGAGGCTGGCGGCGAGGCTCGGGTTCGGAGGACG 130  
Db 609 CGGCGGCGATCGCTGCGGCGCGCGGAGCTGTACGGCTCCAGGTCCTGGCGGACG 550

Qy 131 CGGTGAACCTATGACACAGCGAATTTGCCCGGTTTC 165  
Db 549 CGGTGACGAGACGACCGCGGACGTAACCGGTTTC 515

## RESULT 6

AA459034

LOCUS

DEFINITION

AA459034

VERSION

AA459034.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 423)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs@mail.nih.gov

## FEATURES

source

1. 423  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:6032522"  
/db\_xref="taxon:9606"  
/clone="IMAGE:814365"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GCBI"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCTCATTTTCTTTT-3'  
]. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 5.4%; Score 42.8; DB 9; Length 423;  
Best Local Similarity 48.7%; Pred. No. 39;  
Matches 116; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

Qy 348 CAATCTTCCCGCTCGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAGCGCAGCT 407  
Db 83 CATGTGCTGCTGCGCGGAGCACCGCCCTCCGAGACAGCGGAGCTCCCTACGA 142

Qy 408 CCCACACAGCAATACGACGCGCTGATTTGCTGTTGGGGGGCTGCTTCTGTTC 467  
Db 143 CCTAGCGCTCGCGCCCTCGGGGCCCTATCTCTCTCTGCTGCTGCTTCTC 202

Qy 468 CGCTATGATGCTGGGGGACCTCTGCGGATCTGCTCTCTGCTGCTGCTGCTGCT 527  
Db 203 CATCAGGAGCAGCGCTGCTTCTGCGGCTCGCGGAGCAGCTGCTGCTGCTGCTGCT 262

Qy 528 CTCGCTCGCGGATGAGACGCTGCGAGCTGCAATGCTCAATCTATCCCGGCGAC 585  
Db 263 CACGTCTTCCAGCTGAGCGGAGGCTTTGGGAGGGGCTCTCTGCGCCCCC 320

## RESULT 7

AL513886/c

LOCUS

DEFINITION

AL513886 Homo sapiens PLACENTA

5-PRIME, mRNA sequence.

ACCESSION

AL513886

VERSION

AL513886.2

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

On Feb 13, 2001 this sequence version replaced gi:12777380.

This clone is available royalty-free through LNLML; contact the  
IMAGE Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 382.

Location/Qualifiers  
1. 423

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:6032522"  
/db\_xref="taxon:9606"  
/clone="IMAGE:814365"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GCBI"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCTCATTTTCTTTT-3'  
]. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 5.4%; Score 42.8; DB 9; Length 423;  
Best Local Similarity 48.7%; Pred. No. 39;  
Matches 116; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

Qy 348 CAATCTTCCCGCTCGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAGCGCAGCT 407  
Db 83 CATGTGCTGCTGCGCGGAGCACCGCCCTCCGAGACAGCGGAGCTCCCTACGA 142

Qy 408 CCCACACAGCAATACGACGCGCTGATTTGCTGTTGGGGGGCTGCTTCTGTTC 467  
Db 143 CCTAGCGCTCGCGCCCTCGGGGCCCTATCTCTCTGCTGCTGCTTCTC 202

Qy 468 CGCTATGATGCTGGGGGACCTCTGCGGATCTGCTCTCTGCTGCTGCTGCTGCT 527  
Db 203 CATCAGGAGCAGCGCTGCTTCTGCGGCTCGCGGAGCAGCTGCTGCTGCTGCTGCT 262

Qy 528 CTCGCTCGCGGATGAGACGCTGCGAGCTGCAATGCTCAATCTATCCCGGCGAC 585  
Db 263 CACGTCTTCCAGCTGAGCGGAGGCTTTGGGAGGGGCTCTCTGCGCCCCC 320

## RESULT 7

AL513886/c

LOCUS

DEFINITION

AL513886 Homo sapiens PLACENTA

5-PRIME, mRNA sequence.

ACCESSION

AL513886

VERSION

AL513886.2

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

On Feb 13, 2001 this sequence version replaced gi:12777380.



Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

## FEATURES

source  
Location/Qualifiers  
1..402  
/organism="Chlamydomonas reinhardtii"  
/mol\_type="mRNA"  
/strain="C9"  
/db\_xref="taxon:3055"  
/clone="CM096904\_r"  
/dev\_stage="photoautotrophic growth"  
/clone\_lib="Chlamydomonas reinhardtii C9"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 5.1%; Score 40.6; DB 9; Length 402;  
Best Local Similarity 45.3%; Pred. No. 1.2e+02;  
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;  
QY 42 CGACCTCGTGGGTACATTCGGCTCGTGGCGGCCCTAGGGGGCGCTGCCAGGGCCCT 101  
DB 53 CGAGCTCATCTCGTCAATTTGTGGCGGCACTGCCAATGAGAGCTGCTGAGACCT 112  
QY 102 GCGCATGCGTCCGGTTCTGGAGGACGCGGTGAATATGCAACAGGGAAATTTGCCCGG 161  
DB 113 GCGCGCGCGCGCGAGTGGGAGGGCGCTAGCGCGACAGTCCGTGAGCTTTGGCGCG 172  
QY 162 TTGCTCTTTCTCTATCTTCTCTTGGCTTTGGCTTGTCTGTCTGACCGTTCCAGTTCGCG 221  
DB 173 CCGCAAGGTGTTTGAAGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCAGAACCCAG 232  
QY 222 TTATGAAGTGGCAACGTTGCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAG 281  
DB 233 CTTGCGCGTCCGTGGTGGCCACTCGCTGGGGGGCGGACCGCGGCTGCTGTGAT 292  
QY 282 CATTTGTATGAGCGAGCGGACATGATCATGCAACCCCGGGTGTGCTGCGCTGCTGCTG 341  
DB 293 CTTGATGCACCGACGAGGAGTTTGGCGGCGCATCTACGGGGCGGTGCCCATGCCGG 352  
QY 342 GGGACCAACTCTTCCGCTGCTGGGT 368  
DB 353 CAAGAAGACAGGGCAGCTACATGAT 379

## RESULT 10

AV392165  
LOCUS  
DEFINITION  
AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii  
cDNA clone CM083e05\_r 5', mRNA sequence.  
ACCESSION  
AV392165  
VERSION  
AV392165.1 GI:6546381  
KEYWORDS  
EST.  
SOURCE  
Chlamydomonas reinhardtii  
ORGANISM  
Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadaeae; Chlamydomonas.

## REFERENCE

1 (bases 1 to 551)  
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.  
A large scale structural analysis of cDNAs in a unicellular green  
alga, Chlamydomonas reinhardtii. I. Generation of 3433  
non-redundant expressed sequence tags  
DNA Res. 6 (6), 369-373 (1999)

## JOURNAL

MEDLINE

PUBMED

## COMMENT

Contact: Yasukazu Nakamura  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

## FEATURES

source  
Location/Qualifiers  
1..551  
/organism="Chlamydomonas reinhardtii"  
/mol\_type="mRNA"

/strain="C9"  
/db\_xref="taxon:3055"  
/clone="CM083e05\_r"  
/dev\_stage="photoautotrophic growth"  
/clone\_lib="Chlamydomonas reinhardtii C9"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 5.1%; Score 40.6; DB 9; Length 551;  
Best Local Similarity 45.3%; Pred. No. 1.4e+02;  
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;  
QY 42 CGACCTCGTGGGTACATTCGGCTCGTGGCGGCCCTAGGGGGCGCTGCCAGGGCCCT 101  
DB 108 CGAGCTCATCTCGTCAATTTGTGGCGGCACTGCCAATGAGAGCTGCTGAGGACCT 167  
QY 102 GCGCATGCGTCCGGTTCTGGAGGACGCGGTGAATATGCAACAGGGAAATTTGCCCGG 161  
DB 168 GCGCGCGCGCGCGAGTGGGAGGGCGCTAGCGCGACAGTCCGTGAGCTTTGGCGCG 227  
QY 162 TTGCTCTTTCTCTATCTTCTCTTGGCTTTGGCTTGTCTGTCTGACCGTTCCAGTTCGCG 221  
DB 228 CCGCAAGGTGTTTGAAGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCAGAACCCAG 287  
QY 222 TTATGAAGTGGCAACGTTGCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAG 281  
DB 288 CTTGCGCGTCCGTGGTGGCCACTCGCTGGGGGGCGGACCGCGGCTGCTGTGAT 347  
QY 282 CATTTGTATGAGCGAGCGGACATGATCATGCAACCCCGGGTGTGCTGCGCTGCTGCTG 341  
DB 348 CTTGATGCACCGACGAGGAGTTTGGCGGCGCATCTACGGGGCGGTGCCCATGCCGG 407  
QY 342 GGGACCAACTCTTCCGCTGCTGGGT 368  
DB 408 CAAGAAGACAGGGCAGCTACATGAT 434

## RESULT 11

BI996341  
LOCUS  
DEFINITION  
BI996341  
VERSION  
BI996341.1 GI:16431115  
KEYWORDS  
EST.  
SOURCE  
Chlamydomonas reinhardtii  
ORGANISM  
Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadaeae; Chlamydomonas.

## REFERENCE

1 (bases 1 to 552)  
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,  
Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.  
Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants. Project: 1031  
Unpublished (2001)

## JOURNAL

COMMENT

Contact: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

## FEATURES

source  
1..552  
/organism="Chlamydomonas reinhardtii"  
/mol\_type="mRNA"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Strains II  
(normalized), Lambda Zap II"  
/note="Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:



## COMMENT

Contact: Elizabeth H. Harris  
 DCMB Box 91000  
 Duke University  
 Durham, NC 27708-1000, USA  
 Tel: 919 613 8164  
 Fax: 919 613 8177  
 Email: chlamy@duke.edu.  
 Location/Qualifiers

## FEATURES

source

1. .525  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO<sub>2</sub> and HS medium bubbled with 5% CO<sub>2</sub>. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

## ORIGIN

Query Match 5.0%; Score 40; DB 10; Length 525;  
 Best Local Similarity 47.0%; Pred. No. 1.8e+02;  
 Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 261 GAACGACTCTCCAACTCAAGCATTTGTATGAGCGGCGGACATCATCATCACACCCC 320  
 Db 43 GCACCGCTTCAACAACCCACCCTGCTGAATTTGCGCCGCTCAAGTACCTGTGTCAT 102  
 QY 321 CGGGTGGCTGCCCTGGTTCGGGAGAACAACTCTTCCCGCTGTGGGTAGCGCTACCCC 380  
 Db 103 GGAGGACAGACTCTGCCACTTGAGGTCTATCGATGACTGCGACCTGGGCACTGTCCAA 162  
 QY 381 CACGTCGAGCTAGGAACGCCAGCGTCCCCACGACGACAATACGACGCCACGTCGATTT 440  
 Db 163 CGCGCTCAAAAACAACTCTTCATGATCCCCAACCCCGTCATCGCGCGCGCGCGCGC 222  
 QY 441 GCTCGTTGGGGCGGTCTTTCTGTTCGCTATGATGTTGGGGACCTCTCGCGATCTGT 500  
 Db 223 GGGGACGCGCGCGCGCGCGGAGAGCTAGCGGAGCGCGCGCGGACGCCCATGAAGGT 282  
 QY 501 CTTCCTGCTCTCCAGCTGTTCCAC 524  
 Db 283 CAACATGCGCACCTCTGCTCTAC 306

## RESULT 14

CF846043

LOCUS

DEFINITION pSH036xB09f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sHB Phytophthora sojae cDNA clone sHB036B09 5, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Phytophthora sojae  
 Phytophthora sojae  
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
 Phytophthora.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 534)

Tyler, B.

Not Published

Unpublished (2003)

Contact: Tyler B

Tyler lab

VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA  
 Tel: 540-231-7318

Email: bmtylet@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 036 row: B column: 09

Seq primer: BK reverse primer

High quality sequence stop: 534.

Location/Qualifiers

FEATURES

source

1. .534

/organism="Phytophthora sojae"

/mol\_type="mRNA"

/db\_xref="taxon:67593"

/clone="SHB036B09"

/tissue\_type="mycelium"

/cell\_line="P6497"

/dev\_stage="48 hr. post infection stage"

/lab\_host="Soybean plant"

/clone\_lib="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sHB"

/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 5.0%; Score 40; DB 14; Length 534;  
 Best Local Similarity 46.7%; Pred. No. 1.8e+02;  
 Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 223 TATGAAGTCGCAACGTCGCGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGC 282  
 Db 200 TACGGCGTCGCGGAGATTTACGATATCCGATGCGCATGGCTTCGCGGCTTCTACAACTGGACC 259  
 QY 283 ATTGTGTATGAGCGGCGGACATGATCATGCACACCCCGGGTGGTGGCTCGCTCGG 342  
 Db 260 TCGATGGACCAAGAGAGGCGCGGATCATGCTGACCCCAAGACGGTGGCCAACATTCAC 319  
 QY 343 GAGAACAACTCTTCCCGCTGTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGAAGCC 402  
 Db 320 CACTAGCGGCGCACCATCTCTGCTCGAACCGTGGTGGCTTCGACGTGGACAGATTATC 379  
 QY 403 AGCGTCCCAACCAAGCAATACGACGACGCTGATTTGCTGTTGGGGGGCTGCTTTC 462  
 Db 380 AACTTCTCTGACCAAGAAACGGCTCTCGCAGGTATACGTGATCGCGGTGACGCGCACCCAC 439  
 QY 463 TGTTCGCTATGATGCTGGGGACCTCTGGG 494  
 Db 440 CGTGGCGCAACAGATCTCGGAGAGTCCG 471

## RESULT 15

BI723733

LOCUS

DEFINITION

1031067F08.y1 C. reinhardtii CC-1690, Stress II (normalized),  
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Chlamydomonas reinhardtii  
 Chlamydomonas reinhardtii  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 671)

Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,

Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Sern, D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants. Project: 1031

Unpublished (2001)

Contact: Charles Hauser

DCMB Box 91000

Duke University



Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu.

## FEATURES

source

Location/Qualifiers

1. 671  
/organism="Chlamydomonas reinhardtii"  
/mol\_type="mRNA"  
/strain="CC-1690 wild type mt+ 2lgr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Stress II  
(normalized), Lambda Zap II"  
/note="vector: pBluescript II SK-; Site 1: EcoRI; Site 2:  
XhoI; Stress condition II library, constructed by John  
Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
cells grown to mid-log phase in TAP (NH4+ - containing)  
and shifted to TAP - NO3- (24hrs); H2 production  
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant  
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
PolyA mRNA was purified from each sample, pooled and cDNA  
synthesized. The cDNA was directionally cloned into lambda  
Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')  
sites. pBluescript II SK- plasmids were excised from the  
lambda Zap clones by superinfection with ExAssist  
(Stratagene) phage. The library was normalized using  
method 4 described in Bonaldo et al., (1996) Genome  
Research 6: 791-806."

## ORIGIN

Query Match 5.0%; Score 40; DB 12; Length 671;  
Best Local Similarity 47.0%; Pred. NO. 2e+02;  
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy	261	GAACGACTGCTCCAACTCAAGCATGTGTATGAGCGAGCGGACATGATGCACACCCC	320
Db	353	GCACCGCTTCAACAACCCACCCCGTCTGAATTCCGCCGCTTACAAGTACCTGCTGTCAT	412
Qy	321	CGGGTGGTGGCCCTGCGTTCGGGAGAACAACTCTCCCGCTGCTGGGTAGCGCTCACCCC	380
Db	413	GGAGCAAGACCTGGCCACCTGGAGGTCTCGAGTACTGGACCTGGGCAACCTGTCCAA	472
Qy	381	CACGCTCGCAGCTAGGAACGCGAGCTCCCAACGACGACAAATACGACGCCACGTCGATTT	440
Db	473	CGCGCTCAAAACACACATCTTCATGATCCCAACCCCGTCATCGGGCGCGGGGGGC	532
Qy	441	GCTCGTTGGGGCGGCTGCTTTCTGCTCGCTATGTACGTGGGGGACCTCTGCGGATCTGT	500
Db	533	GGGCGACGGCGGGCGGCGGAGAGTAGCGGAGCGCGCGGCGGAGCCCATGAAGGT	592
Qy	501	CTTCTCTGCTCCAGCTGTTCAC	524
Db	593	CAACATGCGCACCTGCTGCTCAC	616

Search completed: June 16, 2004, 13:27:11  
Job time : 2140.22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 57.3062 Seconds  
(without alignments)  
7698.741 Million cell updates/sec

Title: US-09-899-303A-5

Perfect score: 795

Sequence: 1 ATCTGGTGAAGTCATCGA.....TACTCTTGTCTCCTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795	100.0	795	3	US-08-612-973-5
2	795	100.0	795	3	US-08-927-597-5
3	778.8	98.0	2082	3	US-08-612-973-47
4	778.8	98.0	2082	3	US-08-927-597-47
5	778.8	98.0	2433	3	US-08-612-973-49
6	778.8	98.0	2433	3	US-08-927-597-49
7	703.6	88.5	11076	4	US-09-539-601-1
8	703.6	88.5	11076	4	US-09-539-601-19
9	703.6	88.5	11076	4	US-09-539-601-25
10	703.6	88.5	11076	4	US-09-539-601-31
11	698.8	87.9	1539	2	US-08-470-426B-17
12	698.8	87.9	1863	2	US-08-470-426B-14
13	697.2	87.7	2116	3	US-08-191-160-21
14	695.6	87.5	932	1	US-08-081-072-15
15	695.6	87.5	932	1	US-08-449-093A-15
16	694	87.3	9595	3	US-09-014-416-4
17	694	87.3	9599	3	US-09-014-416-6
18	685.4	86.2	1167	1	US-08-324-977-9
19	685.4	86.2	1167	2	US-08-384-616-9
20	685.4	86.2	1167	2	US-08-904-686A-9
21	685.4	86.2	1167	3	US-09-315-850-9
22	685.4	86.2	1499	1	US-08-324-977-3
23	685.4	86.2	1499	2	US-08-384-616-3
24	685.4	86.2	1499	2	US-08-904-686A-3
25	685.4	86.2	1499	3	US-09-315-850-3
26	685.4	86.2	6039	1	US-08-324-977-11
27	685.4	86.2	6039	2	US-08-384-616-11

28	685.4	86.2	6039	2	US-08-904-686A-11	Sequence 11, Appl
29	685.4	86.2	6039	3	US-09-315-850-11	Sequence 11, Appl
30	685.4	86.2	9030	1	US-08-324-977-13	Sequence 13, Appl
31	685.4	86.2	9030	2	US-08-384-616-13	Sequence 13, Appl
32	685.4	86.2	9030	2	US-08-904-686A-13	Sequence 13, Appl
33	685.4	86.2	9030	3	US-09-315-850-13	Sequence 13, Appl
34	685.4	86.2	9416	1	US-08-324-977-1	Sequence 1, Appl
35	685.4	86.2	9416	2	US-08-384-616-1	Sequence 1, Appl
36	685.4	86.2	9416	2	US-08-904-686A-1	Sequence 1, Appl
37	685.4	86.2	9416	3	US-09-315-850-1	Sequence 1, Appl
38	685.4	86.2	9416	4	US-08-823-895A-27	Sequence 27, Appl
39	684.4	86.1	9472	4	US-08-150-204E-96	Sequence 96, Appl
40	641	80.6	723	3	US-08-612-973-21	Sequence 21, Appl
41	641	80.6	723	3	US-08-927-597-21	Sequence 21, Appl
42	640.8	80.6	742	1	US-08-081-072-18	Sequence 18, Appl
43	640.8	80.6	742	1	US-08-449-093A-18	Sequence 18, Appl
44	628.2	79.0	642	3	US-08-612-973-3	Sequence 3, Appl
45	628.2	79.0	642	3	US-08-927-597-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-08-612-973-5  
; Sequence 5, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 795 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..792  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..789

US-08-612-973-5

Query Match 100.0%; Score 795; DB 3; Length 795;  
Best Local Similarity 100.0%; Pred. NO. 9.8e-201;  
Matches 795; Conservative 0; Mismatches 0; Indels

Qy	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCGCACTTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCGCACTTCGTGGGGTACATT	60
Qy	61	CCGCTCGTGGCGCCGCCCTTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGTCCGGGT	120
Db	61	CCGCTCGTGGCGCCGCCCTTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGTCCGGGT	120
Qy	121	CTGGAGGACGGCGTGAACATAATGCAACAGGGAATTTGCCCGGTGCTTCTTCTATATCTC	180
Db	121	CTGGAGGACGGCGTGAACATAATGCAACAGGGAATTTGCCCGGTGCTTCTTCTATATCTC	180
Qy	181	CTCTTTGGCTTTGCTGCTGTCTGTGACCGGTTCCAGCTTCGGCTATGAAGTGGCAACGTG	240
Db	181	CTCTTTGGCTTTGCTGCTGTGTGACCGGTTCCAGCTTCGGCTATGAAGTGGCAACGTG	240
Qy	241	TCCGGGATGTACCATGTCAAGAAAGACTGCTCCAACTCAAGCATGTGCTATGAGGACAGCG	300
Db	241	TCCGGGATGTACCATGTCAAGAAAGACTGCTCCAACTCAAGCATGTGCTATGAGGACAGCG	300
Qy	301	GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGCTCACCCCAAGCTCGACGTAGGAAAGCCAGCGTCCCAACACAGACA	420
Db	361	TGCTGGGTAGCGCTCACCCCAAGCTCGACGTAGGAAAGCCAGCGTCCCAACACAGACA	420
Qy	421	ATACGACCCACAGCTCGAATTTGCTCTTGGGGCGGCTGCTTTCTGTTCGGCTATGTACGTG	480
Db	421	ATA CGACGCCACGTGCAATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCGGCTATGTACGTG	480
Qy	481	GGGGACCTCTCGGATCTGTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGG	540
Db	481	GGGGACCTCTCGGATCTGTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGG	540
Qy	541	CATGAGCGGTGCGAGACTGCAATTGCTCAATCTATCCGGGCCATAACGGGTCAACGT	600
Db	541	CATGAGCGGTGCGAGACTGCAATTGCTCAATCTATCCGGGCCATAACGGGTCAACGT	600
Qy	601	ATGGCTTGGGATATGATGATGAACCTGGTGTGCTCAACAGGCCCTGGTGGTATCGCAGCTG	660
Db	601	ATGGCTTGGGATATGATGATGAACCTGGTGTGCTCAACAGGCCCTGGTGGTATCGCAGCTG	660
Qy	661	CTCCGGATCCCAAGCTGCTGTGACATGGTGGCGGGGCCCATTTGGGGAGTCTCGGCG	720
Db	661	CTCCGGATCCCAAGCTGCTGTGACATGGTGGCGGGGCCCATTTGGGGAGTCTCGGCG	720
Qy	721	GGTCTTCGCTACTATTTCCATGGTGGGGAACTCGGGCTTAAGGTTTGAATTGTGATCTACTC	780
Db	721	GGTCTTCGCTACTATTTCCATGGTGGGGAACTCGGGCTTAAGGTTTGAATTGTGATCTACTC	780
Qy	781	TTTGCTCCCTAATAG	795
Db	781	TTTGCTCCCTAATAG	795

## RESULT 2

US-08-927-597-5  
: Sequence 5, Application US/08927597  
: Patent No. 6245503  
: GENERAL INFORMATION:  
: APPLICANT: MAERTENS, GEERT  
: APPLICANT: BOSMAN, FONS  
: APPLICANT: DE MARTYNOFF, GUY  
: APPLICANT: BUYSSE, MARIE-ANGE  
: TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPES

QY 361 TGCTGGGTAGCGCTCAGCCCAAGCTCGCAGCTAGGAAGCCAGCGTCCCAACAGACA 420  
Db 361 TGCTGGGTAGCGCTCAGCCCAAGCTCGCAGCTAGGAAGCCAGCGTCCCAACAGACA 420  
QY 421 ATACGACGCGCAGCTCGATTTGCTGCTGGGGGGGCGCTCTTTCTGTTCCGCTATGTAGTG 480  
Db 421 ATACGACGCGCAGCTCGATTTGCTGCTGGGGGGGCGCTCTTTCTGTTCCGCTATGTAGTG 480  
QY 481 GGGGAGCTCTCGCGGATCTGTTCTGCTGCTCCAGCTGTTCAACCATCTCGCTCGCGG 540  
Db 481 GGGGAGCTCTCGCGGATCTGTTCTGCTGCTCCAGCTGTTCAACCATCTCGCTCGCGG 540  
QY 541 CATGACAGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600  
Db 541 CATGACAGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600  
QY 601 ATGGCTTGGGATGATGATGAAGTGGTCCCTACAAAGCGCCCTGGTGGTATCGCAGCTG 660  
Db 601 ATGGCTTGGGATGATGATGAAGTGGTCCCTACAAAGCGCCCTGGTGGTATCGCAGCTG 660  
QY 661 CTCCGGATCCCAAGCTGCTGCTGGGACATGCTGGCGGGGCCCAATGGGGGAGTCTGGCG 720  
Db 661 CTCCGGATCCCAAGCTGCTGCTGGGACATGCTGGCGGGGCCCAATGGGGGAGTCTGGCG 720  
QY 721 GGTCTCGCTACTATTCATGCTGGGGAAGTGGGGAAGTGGGGAAGTGGGGAAGTGGGGA 780  
Db 721 GGTCTCGCTACTATTCATGCTGGGGAAGTGGGGAAGTGGGGAAGTGGGGAAGTGGGGA 780  
QY 781 TTGCTCCCTAATAG 795  
Db 781 TTGCTCCCTAATAG 795

## RESULT 3

US-08-612-973-47  
; Sequence 47, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS B.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2082 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2079  
; FEATURE:  
; NAME/KEY: mat peptide  
; LOCATION: 1..2076  
US-08-612-973-47

Query Match 98.0%; Score 778.8; DB 3; Length 2082;

Best Local Similarity 99.7%; Pred. No. 2.5e-196;

Matches 780; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATCGATACCTTTACATGCGGCTTCCGCGACCTCGTGGGGTACATTCGG 63  
Db 4 TTGGGTAAAGTCAATCGATACCTTTACATGCGGCTTCCGCGACCTCGTGGGGTACATTCGG 63  
QY 64 CTGCTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGTTCGT 123  
Db 64 CTGCTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGTTCGT 123  
QY 124 GAGGAGCGGTGAACATATGCAACAGGGGAATTTGGCCCGGTTGCTCTTTCTATCTTCCTC 183  
Db 124 GAGGAGCGGTGAACATATGCAACAGGGGAATTTGGCCCGGTTGCTCTTTCTATCTTCCTC 183  
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAAGTGTC 243  
Db 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAAGTGTC 243  
QY 244 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGCGAGCGGAC 303  
Db 244 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGCGAGCGGAC 303  
QY 304 ATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
Db 304 ATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
QY 364 TGGGTAGCGCTCACCCCGGCTGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACACCAACAATA 423  
Db 364 TGGGTAGCGCTCACCCCGGCTGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACACCAACAATA 423  
QY 424 CAGCGCCACGCTCGAATTTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 483  
Db 424 CAGCGCCACGCTCGAATTTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 483  
QY 484 GACCTCTGCGGATCTGCTTTCTCTGCTCCAGCTGTTCAACCATCTCGCTCGCGCGCAT 543  
Db 484 GACCTCTGCGGATCTGCTTTCTCTGCTCCAGCTGTTCAACCATCTCGCTCGCGCGCAT 543  
QY 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACACATAAGCGGTCAACGTATG 603  
Db 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACACATAAGCGGTCAACGTATG 603  
QY 604 GCTTGGGATATGATGATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 663  
Db 604 GCTTGGGATATGATGATGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 663  
QY 664 CGGATCCCAACAGCTGCTGCTGGAATGCTGGGCGGGGGCCCAATGGGGGAGTCTTGGCGGGT 723  
Db 664 CGGATCCCAACAGCTGCTGCTGGAATGCTGGGCGGGGGCCCAATGGGGGAGTCTTGGCGGGT 723  
QY 724 CTGCGCTACTATTTCCATGGTGGGGAACCTGGGCTTAAGGTTTGTGATTTGATGCTACTCTTT 783  
Db 724 CTGCGCTACTATTTCCATGGTGGGGAACCTGGGCTTAAGGTTTGTGATTTGATGCTACTCTTT 783  
QY 784 GC 785  
Db 784 GC 785

```
RESULT 4
US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47

Query Match      98.0%; Score 778.8; DB 3; Length 2082;
Best Local Similarity 99.7%; Pred. No. 2.5e-196;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 TTGGGTAAAGGTGATGATACCTTACATCGGGCTTCGCCGACTCGTGGGGTACATTCGG 63
DB      4 TTGGGTAAAGGTGATGATACCTTACATCGGGCTTCGCCGACTCGTGGGGTACATTCGG 63

QY      64 CTCGTGGGGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCTCGGGTTCG 123
DB      64 CTCGTGGGGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCTCGGGTTCG 123

QY      124 GAGGACGGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
DB      124 GAGGACGGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183

QY      184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
DB      184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

US-08-927-597-47
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47

US-08-927-597-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION:
; NAME: BYRNE, THOMAS E.
```



```
QY 64 CTGCTGGGCGCCCTTACAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCCGGGTTCTG 123
Db 415 CTGCTGGGCGCCCTTACAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGGTGCTCTTTCTATCTCTCTC 183
Db 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGGTGCTCTTTCTATCTCTCTC 534
QY 184 TTGGCTTTTCTGCTGCTGCTGAGCGTTCCAGCTTCCGCTTATGAAGTGGCCAAAGCTGTC 243
Db 535 TTGGCTTTTCTGCTGCTGCTGAGCGTTCCAGCTTCCGCTTATGAAGTGGCCAAAGCTGTC 594
QY 244 GGGATGTACCATGTCAACAGGAGCTGCTCCAACTCAAGCAATTTGTATGAGGAGCGGAC 303
Db 595 GGGATGTACCATGTCAACAGGAGCTGCTCCAACTCAAGCAATTTGTATGAGGAGCGGAC 654
QY 304 ATGATCATGACACCCCGGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 655 ATGATCATGACACCCCGGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
QY 364 TGGGTAGCGCTCACCCCGCGCTCGCAGCTAGGAAAGCGCAGAGTCCCGCACGCAATA 423
Db 715 TGGGTAGCGCTCACCCCGCGCTCGCAGCTAGGAAAGCGCAGAGTCCCGCACGCAATA 774
QY 424 CGACGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 775 CGACGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 484 GACCTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 835 GACCTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 544 GAGACGCTGACGAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Db 895 GAGACGCTGACGAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
QY 664 CGGATCCCAAGAGCTGCTGACATGCTGGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 723
Db 1015 CGGATCCCAAGAGCTGCTGACATGCTGGGCGGGGGCCCATTTGGGAGTCTTGGCGGG 1074
QY 724 CTGCGCTACTATTCCATGCTGGGAACTGGGCTTAAGGTTTGAATGCTGATGCTACTCTT 783
Db 1075 CTGCGCTACTATTCCATGCTGGGAACTGGGCTTAAGGTTTGAATGCTGATGCTACTCT 1134
QY 784 GC 785
Db 1135 GC 1136
```

```
RESULT 7
US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Barten Schlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
```

```
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B; parental sequence
; OTHER INFORMATION: without cell culture-adaptive mutations
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-1

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGAGCTTCGTTGGGTACATTTCCG 63
Db 2167 TTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGAGCTTCGTTGGGTACATTTCCG 2226
QY 64 CTGCTGGGCGCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCGTTGGT 123
Db 2227 CTGCTGGGCGCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCGTTGGT 2286
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGGTGCTCTTTCTATCTCTCTC 183
Db 2287 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGGTGCTCTTTCTATCTCTCTC 2346
QY 184 TTGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
Db 2347 TTGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
QY 244 GGGATGACCATGTCAAGAAAGCTGCTCAACTCAAGCAATTTGTATGAGGAGCGGAC 303
Db 2407 GGGATGACCATGTCAAGAAAGCTGCTCAACTCAAGCAATTTGTATGAGGAGCGGAC 2466
QY 304 ATGATCATGACACCCCGGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 2467 ATGATCATGACACCCCGGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526
QY 364 TGGGTAGCGCTCACCCCGCGCTGCGAGCTTAGGAAAGCGCAGCGTCCCGCACGCAATA 423
Db 2527 TGGGTAGCGCTCACCCCGCGCTGCGAGCTTAGGAAAGCGCAGCGTCCCGCACGCAATA 2586
QY 424 CGACGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 2587 CGACGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2646
QY 484 GACCTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 2647 GACCTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2706
QY 544 GAGACGGGTGAGGAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Db 2707 GAGACGGGTGAGGAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2766
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 2767 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2826
QY 664 CGGATCCCAAGAGCTGCTGACATGCTGGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 723
```



```

Db      2827 CGGATCCCAAGCTGTGTGATATATGTTGGGGGGGCCCCATTTGGGGAGTCTTAGCGGGC 2886
QY      724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTGTGATGCTACTCTTT 783
      |||||||
Db      2887 CTTCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATTTGTGATGCTACTCTTT 2946
QY      784 GC 785
      ||
Db      2947 GC 2948

RESULT 8
US-09-539-601-19
; Sequence 19, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/9-13P
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell
; OTHER INFORMATION: culture-adaptive mutations from clone 9-13P
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-19

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY      4 TTGGGTAAGTCTATCGATACCTTACATGGCGCTTCGCGGACCTCGTGGGGTACATTCGG 63
      |||||||
Db      2167 TTGGGTAAGTCTATCGATACCTTACATGGCGCTTCGCGGACCTCGTGGGGTACATTCGG 2226
QY      64 CTGCTGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGCTCGGTTCTG 123
      |||||||
Db      2227 CTGCTGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGCTCGGTTCTG 2286
QY      124 GAGGACGGCTGAACTATGCAACGGGAATTTGCCCGGTGCTTTCTCTATCTTCTC 183
      |||||||
Db      2287 GAGGACGGCTGAACTATGCAACGGGAATTTGCCCGGTGCTTTCTCTATCTTCTC 2346
QY      184 TTGCGTTTGTCTGCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGCTGTCC 243
      |||||||
Db      2347 TTGCGTTTGTCTGCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGCTATCC 2406

```

```

QY      244 GGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCGGAC 303
      |||||||
Db      2407 GGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCGGAC 2466
QY      304 ATGATCATGACACCCCGGGTGGTCCCTTGGGTTGGGAGAACAACTTCCCGCTGC 363
      |||||||
Db      2467 ATGATCATGACATCCCGGGTGGTCCCTTGGGTTGGGAGAACAACTTCCCGCTGC 2526
QY      364 TGGGTAGCGCTCAGCCCGACGCTCGACGTAGGAAACGACGCTCCCGACACGACAAATA 423
      |||||||
Db      2527 TGGGTAGCGCTCAGCTCCACGCTCGGACGAGGACGCTAGCGTCCCGACCTACGAGATA 2586
QY      424 CGACGCCACGCTCGAATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTTGGG 483
      |||||||
Db      2587 CGACGCCATGTTCGATTTGCTGTTGGGGCGGCTGCTTTCTGCTCTCGCTATGTACGTTGGG 2646
QY      484 GACCTCTGGGATCTGCTTCTCTGCTCCAGCTGTTCCACCATCTCGCTCGCGCGGAT 543
      |||||||
Db      2647 GATCTCTGGGATCTGTTTCTGCTGCTCGCGGCTGTTTCACTTCTCGCTCGCGGAC 2706
QY      544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAAGGTTCAAGTATG 603
      |||||||
Db      2707 GAGACGATACGAGGACTGCAATTTGCTCAATATATCCCGGCCACGTCAGGTCAAGTATG 2766
QY      604 GCTTGGGATATGATGATGAATGCTGCTGCTTACAAACGCGGCTGCTGCTATGCGAGTCTC 663
      |||||||
Db      2767 GCTTGGGATATGATGATGAATGCTGCTGCTTACAAACGCGGCTGCTGCTATGCGAGTCTC 2826
QY      664 CGGATCCCAAGCTGTCTGACATGTTGGCGGGGCCCATTTGGGGAGTCTCTGCGGGT 723
      |||||||
Db      2827 CGGATCCCAAGCTGTCTGATATGTTGGCGGGGCCCATTTGGGGAGTCTCTGCGGGC 2886
QY      724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTCATGCTACTCTTT 783
      |||||||
Db      2887 CTTCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTCATGCTACTCTTT 2946
QY      784 GC 785
      ||
Db      2947 GC 2948

RESULT 9
US-09-539-601-25
; Sequence 25, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/5.1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from

```

```

; OTHER INFORMATION: encephalomyocarditis virus
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell
; OTHER INFORMATION: culture-adaptive mutations of clone 5.1
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-25

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 4 TTGGTAAAGTTCATCGATACCTTACATCGCGCTTGGCGGACCTCGTGGGTACATTCGG 63
Db 2167 TTGGTAAAGTTCATCGATACCTTACATCGCGCTTGGCGGACCTCGTGGGTACATTCGG 2226

Qy 64 CTGCTGGCGCCCTTGGGCGGCGCTGCCAGGCGCTTGGGCGATGGCGTTCGGGTTCG 123
Db 2227 CTGCTGGCGCCCTTGGGCGGCGCTGCCAGGCGCTTGGGCGATGGCGTTCGGGTTCG 2286

Qy 124 GAGGACGCGTGAATCATGCAACAGGAAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
Db 2287 GAGGACGCGTGAATCATGCAACAGGAAATTTGCCCGGTGCTCTTTCTATCTTCCTC 2346

Qy 184 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 243
Db 2347 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 2406

Qy 244 GGGATGTACCATGTCACCAAGACGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 303
Db 2407 GGGATGTACCATGTCACCAAGACGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 2466

Qy 304 ATGATCATGCAACCCCGGGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 2467 ATGATCATGCAACCCCGGGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526

Qy 364 TGGGTAGGCTTACCCCGGCGTGGCGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 423
Db 2527 TGGGTAGGCTTACCCCGGCGTGGCGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2586

Qy 424 CGAGCCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 2587 CGAGCCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2646

Qy 484 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
Db 2647 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2706

Qy 544 GAGAGGTGCGAGGCTGCAATTCATCTATCTATCTATCTATCTATCTATCTATCTATCT 603
Db 2707 GAGAGGTGCGAGGCTGCAATTCATCTATCTATCTATCTATCTATCTATCTATCTATCT 2766

Qy 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
Db 2767 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2826

Qy 664 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
Db 2827 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2886

Qy 724 CTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTGAATGATGATGATGATGATGAT 783
Db 2887 CTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTGAATGATGATGATGATGATGAT 2946

Qy 784 GC 785
Db 2947 GC 2948

```

```

RESULT 10
US-09-539-601-31
; Sequence 31, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 31
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct 1389/Core-3'/19
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell culture
; OTHER INFORMATION: adaptive mutations from clone no. 19
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-31

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 4 TTGGTAAAGTTCATCGATACCTTACATCGCGCTTGGCGGACCTCGTGGGTACATTCGG 63
Db 2167 TTGGTAAAGTTCATCGATACCTTACATCGCGCTTGGCGGACCTCGTGGGTACATTCGG 2226

Qy 64 CTGCTGGCGCCCTTGGGCGGCGCTGCCAGGCGCTTGGGCGATGGCGTTCGGGTTCG 123
Db 2227 CTGCTGGCGCCCTTGGGCGGCGCTGCCAGGCGCTTGGGCGATGGCGTTCGGGTTCG 2286

Qy 124 GAGGACGCGTGAATCATGCAACAGGAAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
Db 2287 GAGGACGCGTGAATCATGCAACAGGAAATTTGCCCGGTGCTCTTTCTATCTTCCTC 2346

Qy 184 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 243
Db 2347 TTGGCTTTGCTGCTGCTGACCGTTCCAGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 2406

Qy 244 GGGATGTACCATGTCACCAAGACGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 303
Db 2407 GGGATGTACCATGTCACCAAGACGCTTCCAGCTTGTGTATGAAGTGGCGACGTCC 2466

Qy 304 ATGATCATGCAACCCCGGGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 2467 ATGATCATGCAACCCCGGGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526

Qy 364 TGGGTAGGCTTACCCCGGCGTGGCGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 423
Db 2527 TGGGTAGGCTTACCCCGGCGTGGCGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2586

Qy 424 CGAGCCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 2587 CGAGCCAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2646

Qy 484 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
Db 2647 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2706

Qy 544 GAGAGGTGCGAGGCTGCAATTCATCTATCTATCTATCTATCTATCTATCTATCTATCT 603
Db 2707 GAGAGGTGCGAGGCTGCAATTCATCTATCTATCTATCTATCTATCTATCTATCTATCT 2766

Qy 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
Db 2767 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2826

Qy 664 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
Db 2827 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2886

Qy 724 CTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTGAATGATGATGATGATGATGAT 783
Db 2887 CTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTGAATGATGATGATGATGATGAT 2946

Qy 784 GC 785
Db 2947 GC 2948

```

Db 2527 TGGTAGCGCTCACTCCACGCTCGCGCCAGGAACGCTAGCGTCCCGCCACATACGACGATA 2586  
Qy 424 CGAGCCACAGCTCGATTGCTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGATGGG 483  
Db 2587 CGAGCCACATGCTGATTGCTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGATGGG 2646  
Qy 484 GACCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
Db 2647 GATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2706  
Qy 544 GAGACGCTGAGGAGTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCT 603  
Db 2707 GAGACGCTGAGGAGTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCT 2766  
Qy 604 GCTTGGGAT 663  
Db 2767 GCTTGGGAT 2826  
Qy 664 CGATCCCAAGCTGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723  
Db 2827 CGATCCCAAGCTGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2886  
Qy 724 CTGCGCTACTATCCATGCTGGGAACTGGGCTAAGGTTTGGATGATGATGATGATGATGATGATGATGAT 783  
Db 2887 CTGCGCTACTATCCATGCTGGGAACTGGGCTAAGGTTTGGATGATGATGATGATGATGATGATGATGAT 2946  
Qy 784 GC 785  
Db 2947 GC 2948

## RESULT 11

US-08-470-426B-17  
; Sequence 17, Application US/08470426B  
; Patent No. 5856458  
; GENERAL INFORMATION:  
; APPLICANT: Okamoto, Hiroaki  
; APPLICANT: Nakamura, Tetsuo  
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR  
; TITLE OF INVENTION: APPLICATION FOR HIGH-FAIDELITY DETECTION OF NON-A, NON-B  
; TITLE OF INVENTION: HEPATITIS VIRUS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,  
; STREET: 1850 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,426B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-153402  
; FILING DATE: 12-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weillacher, Robert G.  
; REGISTRATION NUMBER: 20,531  
; REFERENCE/DOCKET NUMBER: 06/59-47083.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 659-2811  
; TELEFAX: (202) 659-1462  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1539 base pairs  
; TYPE: nucleic acid

; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
US-08-470-426B-17  
Query Match 87.9%; Score 698.8; DB 2; Length 1539;  
Best Local Similarity 93.4%; Pred. No. 3.1e-175;  
Matches 730; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
Qy 4 TTGGTAAAGTCATCGATACCTTACATGCGGCTTTCGCGGACCTCGTGGGGTACATTCG 63  
Db 355 TTGGTAAAGTCATCGATACCTTACATGCGGCTTTCGCGGACCTCGTGGGGTACATTCG 414  
Qy 64 CTGCTGCGG 123  
Db 415 CTGCTGCGG 474  
Qy 124 GAGNACGCGG 183  
Db 475 GAGNACGCGG 534  
Qy 184 TTGGCTTTGCT 243  
Db 535 TTGGCTTTGCT 594  
Qy 244 GGGATGTACCATGTGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGCGACGGAC 303  
Db 595 GGGATGTACCATGTGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGCGACGGAC 654  
Qy 304 ATGATCATGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 363  
Db 655 ATGATCATGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 714  
Qy 364 TGGGTAGCGCTCACCGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
Db 715 TGGGTAGCGCTCACCGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774  
Qy 424 CGACGCCACGCTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
Db 775 CGACGCCACGCTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834  
Qy 484 GACCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
Db 835 GATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894  
Qy 544 GAGACGCTGACGAGTCAATTTGCTCAATCTATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603  
Db 895 GAGACGCTGACGAGTCAATTTGCTCAATCTATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 954  
Qy 604 GCTTGGGAT 663  
Db 955 GCTTGGGAT 1014  
Qy 664 CGGATCCCAAGCTGCTGCGGACATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723  
Db 1015 CGGATCCCAAGCTGCTGCGGACATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1074  
Qy 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGGATGATGATGATGATGATGATGATGATGATGAT 783  
Db 1075 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGGATGATGATGATGATGATGATGATGATGATGAT 1134  
Qy 784 GC 785  
Db 1135 GC 1136

## RESULT 12

US-08-470-426B-14  
; Sequence 14, Application US/08470426B  
; Patent No. 5856458  
; GENERAL INFORMATION:  
; APPLICANT: Okamoto, Hiroaki  
; APPLICANT: Nakamura, Tetsuo



INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2116 base pairs  
TYPE: nucleotide with corresponding protein  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
ORIGINAL SOURCE:  
ORGANISM: human; serum infectious for PT-NANBH  
IMMEDIATE SOURCE:  
LIBRARY: cDNA clones from 5' end of the genome  
FEATURE:  
LOCATION: from 308 to 2116 bp start of the PT-NANBH  
LOCATION: polyprotein  
OTHER INFORMATION: viral structural and non-structural  
OTHER INFORMATION: proteins  
US-08-191-160-21

Query Match 87.7%; Score 697.2; DB 3; Length 2116;  
Best Local Similarity 93.2%; Pred. No. 9e-175;  
Matches 729; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	4	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	63
DB	662	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	721
QY	64	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123
DB	722	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	781
QY	124	GAGGACGGGTGAACATATGACAAAGGGAATTTCCCGGTTCCTTTCTATCTTCCTC	183
DB	782	GAGGACGGGTGAACATATGACAAAGGGAATTTACCCGGTTGCTTTCTATCTTCCTC	841
QY	184	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAGTGGCGAAGTTC	243
DB	842	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAGTGGCGAAGTTC	901
QY	244	GGGATGATACATGTCACGAAGCACTCTCCAACTCAAGCAATGCTATGAGGAGCGGAC	303
DB	902	GGGATGATACATGTCACGAAGCAATGCTCCAACTCAAGCAATGCTATGAGGAGCGGAC	961
QY	304	ATGATCATGCACACCCCGGGTGCCTTCGCTTCGGGAGAACAACTCTTCCCGCTGC	363
DB	962	ATGATCATGCACACCCCGGGTGCCTTCGCTTCGGGAGGAATTAATCTCCCGCTGC	1021
QY	364	TGGGTAGCCCTACCCCGACGCTCGGAGTAGGAACCGCGCTCCCGACAGACAATA	423
DB	1022	TGGGTAGCCCTACCTCCACGCTCGCGGCGCAAGGAGCCAGCATCCCGACAGACAATA	1081
QY	424	CGAGCGCAGTTCGATTTGCTTCGTTGGGGGGCTGCTTCTGTTTCGCTATGACGTGGG	483
DB	1082	CGAGCGCAGTTCGATTTGCTTCGTTGGGGGGCTGCTTCTGTTTCGCTATGACGTGGG	1141
QY	484	GACCTCTGGGATCTCTTCTCTCCAGCTGTTTCAACATCTCGCTCCCGCGGCAAT	543
DB	1142	GATCTCTGGGATCTCTTCTCTCCAGCTGTTTCAACATCTCTCGCTCCCGCGGCAAT	1201
QY	544	GAGAGCGTGCAGGATGCTCAATTCATATCCCGGCGCACATAACGGGTACCGTATG	603
DB	1202	CAGAGCGTGCAGGATGCTCAATTCATATCCCGGCGCACATATCAGGTACCGCATG	1261
QY	604	GCTTGGGATATGATGATGATGCTGCTCCCTACAAAGCGCTGGTGTATCGGACGTCTC	663
DB	1262	GCTTGGGATATGATGATGATGCTGCTCCCTACAAAGCGCTGGTGTATCGGACGTCTC	1321
QY	664	CGGATCCCAAGCTGCTGTGGACATGGTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT	723
DB	1322	CGGATCCCAAGCTGCTGTGGACATGGTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGG	1381
QY	724	CTCGCTACTATTTCCATGTTGGGGAATCTGGGCTTAAGTTTGTATGTGATCTCTTT	783
DB	1382	CTTGCCTACTATTTCCATGTTGGGGAATCTGGGCTTAAGTTTGTATGTGATCTCTTT	1441

QY 784 GC 785  
DB 1442 GC 1443

## RESULT 14

US-08-081-072-15  
Sequence 15, Application US/08081072  
Patent No. 5641654  
GENERAL INFORMATION:  
APPLICANT: No. 5641654oru MAKI, Kenjiro YAMAGUCHI, Ayumi  
TITLE OF INVENTION: TOYOSHIMA, and Michinori KOHARA  
TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC  
TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS  
TITLE OF INVENTION: DIAGNOSIS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman  
STREET: 130 Water Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-4280  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk, 3.50inch, 1.4Mb storage  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh Classic  
SOFTWARE: Microsoft Word Version 4.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/081,072  
FILING DATE: June 22, 1993  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 932 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
US-08-081-072-15

Query Match 87.5%; Score 695.6; DB 1; Length 932;  
Best Local Similarity 93.1%; Pred. No. 1.8e-174;  
Matches 728; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY	4	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	63
DB	9	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	68
QY	64	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGTTCG	123
DB	69	CTTGTGCGGGCCCCCTTAGGGGGTGTCTGCAGGGCCCTGGCACATGGTGTCCGGTTCG	128
QY	124	GAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGTTCCTTTCTATCTTCCTC	183
DB	129	GAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGTTCCTTTCTATCTTCCTC	188
QY	184	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAGTGGCGAAGTTC	243
DB	189	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAGTGGCGAAGTTC	248
QY	244	GGGATGATACATGTCACGAAGCACTCTCCAACTCAAGCAATGCTATGAGGAGCGGAC	303
DB	249	GGGATGATACATGTCACGAAGCACTCTCCAACTCAAGCAATGCTATGAGGAGCGGAC	308
QY	304	ATGATCATGCACACCCCGGGTGCCTTCGCTTCGGGAGAACAACTCTTCCCGCTGC	363
DB	309	ATGATCATGCATACCCCGGGTGCCTTCGCTTCGGGAGAACAACTCTTCCCGCTGC	368
QY	364	TGGGTAGCCCTACCCCGACGCTCGGAGTAGGAACCGCGCTCCCGACAGACAATA	423
DB	369	TGGGTAGCCCTACCTCCACGCTTAGCGGCGCAGGAACACCGAGCTCCCGACAGACAATA	428



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1473.55 Seconds  
(without alignments)  
14206.949 Million cell updates/sec

Title: US-09-899-303A-9  
Perfect score: 483  
Sequence: 1 ATGCCCGGTGCTCTTCTC.....TGATGAACCTGGTCTAATAG 483

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBml:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_stb.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_stb.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	483	100.0	483	6	A48671	A48671 Sequence 9
2	483	100.0	483	6	AR157327	AR157327 Sequence
3	483	100.0	483	6	AX452758	AX452758 Sequence
4	483	100.0	483	6	AX685010	AX685010 Sequence
5	483	100.0	483	6	AX794852	AX794852 Sequence
6	474.2	98.2	480	6	A48673	A48673 Sequence 11
7	474.2	98.2	480	6	AR157328	AR157328 Sequence
8	474.2	98.2	480	6	AX452760	AX452760 Sequence
9	474.2	98.2	480	6	AX685012	AX685012 Sequence
10	474.2	98.2	480	6	AX794854	AX794854 Sequence
11	438.2	90.7	9379	14	AF207766	AF207766 Hepatitis C
12	435	90.1	1880	14	HPC5TRJ4	D00832 Hepatitis C
13	435	90.1	2540	6	E04260	E04260 cDNA encodi
14	435	90.1	2540	6	E04805	E04805 cDNA to 5'
15	435	90.1	2540	6	E07391	E07391 cDNA encodi
16	433.4	89.7	1539	6	AR027786	AR027786 Hepatitis C
17	433.4	89.7	1539	6	AR027783	AR027783 Sequence
18	433	89.6	9580	14	AF054250	AF054250 Hepatitis
19	433	89.6	9580	14	AF054257	AF054257 Hepatitis
20	431.8	89.4	8780	14	AF054258	AF054258 Hepatitis
21	431.8	89.4	8780	14	AF054255	AF054255 Hepatitis
22	430.2	89.1	8781	14	AF054254	AF054254 Hepatitis
23	430.2	89.1	8781	14	AF165052	AF165052 Hepatitis
24	430.2	89.1	8781	14	AF165052	AF165052 Hepatitis
25	430.2	89.1	8781	14	HPCJ491	D10750 Hepatitis C
26	430.2	89.1	9595	6	AR119832	AR119832 Sequence
27	430.2	89.1	9595	14	AF054247	AF054247 Hepatitis
28	430.2	89.1	9596	14	AF054249	AF054249 Hepatitis
29	430.2	89.1	9599	6	AR119833	AR119833 Sequence
30	428.6	88.7	577	6	E04085	E04085 gDNA encodi
31	428.6	88.7	618	14	HPCB1P12	D28929 Hepatitis C
32	428.6	88.7	8779	14	AF054251	AF054251 Hepatitis
33	428.6	88.7	8780	14	AF054252	AF054252 Hepatitis
34	428.6	88.7	9377	14	AF207756	AF207756 Hepatitis
35	428.6	88.7	9410	14	HPCK1R2	D50481 Hepatitis C
36	428.6	88.7	9435	14	AB049093	AB049093 Hepatitis
37	428.6	88.7	9595	14	AF054248	AF054248 Hepatitis
38	427	88.4	1595	14	HPCNS1SPF	M74809 Hepatitis C
39	427	88.4	1635	14	HPCNS1SPD	M74807 Hepatitis C
40	427	88.4	3296	14	AB008446	AB008446 Hepatitis
41	427	88.4	8780	14	AF054259	AF054259 Hepatitis
42	427	88.4	8781	14	AF054256	AF054256 Hepatitis
43	427	88.4	9361	14	AF483269	AF483269 Hepatitis
44	427	88.4	9369	14	AF165054	AF165054 Hepatitis
45	427	88.4	9379	14	AF165051	AF165051 Hepatitis

ALIGNMENTS

RESULT 1						
A48671	A48671	Sequence 9 from Patent WO9604385.	483 bp	DNA	linear	PAT 07-MAR-1997
LOCUS						
DEFINITION	A48671	Sequence 9 from Patent WO9604385.				
ACCESSION	A48671					
VERSION	A48671.1	GI:2302384				
KEYWORDS						
SOURCE		unidentified				
ORGANISM		unidentified				
REFERENCE		1 (bases 1 to 483)				
AUTHORS		Maertens,G., Bosman,F., De,M.G. and Buyse,M.				
TITLE		PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE				
JOURNAL		Patent: WO 9604385-A 9 15-FEB-1996;				













```

/protein_id="AAF65956.1"
/db_xref="GI:7650250"

```

	162	GGAAGCCGACGGTCCCACAAACGACAATACGACGCCACCGTCGATTGTCTGTTGGGGCTG	301
Df	1015	GCCTTGCGGGAGAACAACACTCCTCCCGTTGCTGGGTGGCGCTCACTCCACACGCTCGCGGCCA	1074
Qy	242	GGAAGCCGACGGTCCCACAAACGACAATACGACGCCACCGTCGATTGTCTGTTGGGGCTG	301
Df	1075	GGAAGCCGACGGTCCCACCACTTAAGACAATACGACGCTCAAGGTCGATTGTCTGTTGGGGCGG	1134
Qy	302	CTGCTTTCTGTTGGCTATGTAGTCGTGGGGATCTCTCGGATCTGTTTCTGTTGTTCC	361
Df	1135	CTGCTTTCTGTTGGCTATGTAGTCGTGGGGATCTCTCGGATCTGTTTCTGTTGTTCC	1194
Qy	362	AGCTGTTTCACCTTCTCACCTCGCCGGCGCATCAAACAGTACAGGACTGCAACTGCTCAATCT	421
Df	1195	AGCTGTTTCACCTTCTGCTTGGCGGATGAGACGGTACAGGACTGCAATGCTCAATCT	1254
Qy	422	ATCCCCGCCCATGTATCAGGTCACC CGCATGCTTTGGGATATGATGATGAAC TGTC	476
Df	1255	ATCCCCGCCCATGTATCAGGTCACCGCATGCTTTGGGATATGATGATGAAC TGTC	1309
RESULT 12			
HPCSTRJ4			
LOCUS			
DEFINITION	HPCSTRJ4	1880 bp RNA linear VRL 16-MAY-1998	
	Hepatitis C virus gene for structural protein, partial cds,		
	isolate:HC-J4.		
ACCESSION	D00832		
VERSION	D00832.1 GI:221513		
KEYWORDS	structural protein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
	Hepadnavirus.		
REFERENCE	1 (bases 1 to 1880)		
AUTHORS	Okamoto,H., Okada,S., Sugiyama,Y., Yotsumoto,S., Tanaka,T., Yoshizawa,H., Tsuda,F., Miyakawa,Y. and Mayumi,M.		
TITLE	The 5'-terminal sequence of the hepatitis C virus genome		
JOURNAL	Jpn. J. Exp. Med. 60 (3), 167-177 (1990)		
MEDLINE	91013116		
PUBMED	2170712		
REFERENCE	2 (sites)		
AUTHORS	Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka.H., Machida,A., Miyakawa,Y. and Mayumi,M.		
TITLE	Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions		
JOURNAL	J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)		
MEDLINE	92044440		
PUBMED	1658196		
REFERENCE	3 (sites)		
AUTHORS	Mink,M.A., Benichou,S., Madaule,P., Tiollais,P., Prince,A.M. and Inchauspe,G.		
TITLE	Characterization and mapping of a B-cell immunogenic domain in Hepatitis C virus E2 glycoprotein using a yeast peptide library		
JOURNAL	Virology 200 (1), 246-255 (1994)		
MEDLINE	94174722		
PUBMED	7510436		
REFERENCE	4 (sites)		
AUTHORS	Hotta,H., Doi,H., Hayashi,T., Purwanta,M., Soemarto,W., Mizokami,M., Ohba,K. and Homma,M.		
TITLE	Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia		
JOURNAL	Arch. Virol. 136 (1-2), 53-62 (1994)		
MEDLINE	94270990		
PUBMED	7545932		
COMMENT	These data kindly submitted in computer readable form by: Hiroaki Okamoto Immunology Division, Jichi Medical School Kawachi-gun, Tochigi-ken 329-04 Japan Phone: 0285-44-2111 x3334 Fax: 0285-44-1557. Location/Qualifiers		
FEATURES			

/t=translation="MSTNPKPQRKTRNNRRPODVKPPGGQIVGVGVLLPBRGPRL  
 GVARTRKTSRSPRRGRQRP1PKARKPSPRTWAGYQVFWPLYVNGSNGWAGWLLSPRG  
 SRWSWGTDPRRSRNGKIDTUTCPGADLMGYTLPVAGVPGAGARALAGWLLSPVLE  
 GVNATYKPCFCSF1FLLALLSCLTTPASAVENRANSGVTHVLLDNCNSVTPANAD  
 MN1MHTGCVFCPSRENNSSRWVALLPTLAAARNASVPTKTRTHVLLDVLGAAAFCSAM  
 VGDLCGSVVLVSQVTPSRHETVCDCNCS1PCHLSKGLHMAWMMNMSPTPALV  
 SOLLR1PQVAMDVAGAHGVLVGLAGYVSMVGNKALH1VLLLPAGVDSGTSSVGVTA  
 GRITVGLARPTSPSQNIQLVNTNGSHIINRTALCNDLSLQTF1AALPYTHYKND  
 GCPERMASCRS1DKFPGWGP1VTAGDSHLSRDPYCWHPAGVPCG1PAAQVCGVP1C  
 TSSPVVGTSDRSRGVTPNGVNSVLLDNTPRCPWNPQPCGTPMMSNTGFTKCGG  
 TILKVMGVGNVNTL1CPTDKFRKPEATYTKCGSGPWLTPCVDYVPTRLWHPYPCVTFN  
 PTK1KVMGVGGEHRLNAAACNTEGRCENCLBNDRSELSPLALLSTEMOVLPCSPFTL  
 PALSTGL1HLHQNVVDQVLYXIGISAVSVPIKWEVLY1PILLADARVAC1LWMLLL  
 L1QAQBALEK1V1LNAASVAGANGHLSFLVLPFCANVY1KGLV1KGLVNAAY1XGVWPLLL  
 LLLAP1PLAY1KEMASVAGAVFVGLVLL1TSPHYKVF1LARL1WMLQ1L1TRAEAL  
 LQWVP1NTP1NGRDA1ILLTCAAHPEL1POISKAL1LGP1VL1OAGITRVPVPR  
 AQGL1RYMLVRYKAGVGHQWALMLAALQY1TVDHL1PTLWDHAHGLD1LAVALVE  
 P1VFSDBMETK1TWGAT2ACNWI1SGLPV5ARRRKE1LLGPAVDSHGRRWLLAP1T  
 AYCOQTPLEG1LC1TSLTRDKNOVGEVOVUSTAQSP1LACVAGVCHVWVYHGAGSK  
 TLAPKGP1QTYM1NVQDL1VGL1AP1PGARSL1PCTCGSSDL1LTRHADV1PVRKRG  
 DRSGLTNSPRV1Y1KGSQSGFL1CPSGHVGI1FRAAV1TRGVAKAV1PVPVSMETL  
 TRSPVFNDS1PPAVQTP1VQAHHLHAGT1CGSKGTSPAPAA1YAKV1VLNFSVAATL  
 GFGAMSKAGHVDNI1RTG1T1VGA1T1PSTYKFLDAGCGSGAVD1IMCDBEYK  
 TDGTS1LIG1TGLVQDA1TGAAR1V1LAP1TSPG1VTPHENT1E1AL1SNTG1PYGK  
 A1PI1BJ1KGRH1L1PCHSKKCD1ELAK1SLGL1NAY1TGLVDSV1PVSVDGNV1G  
 TDLAM1GFT1GDLDS1DNCVT1QV1DPS1LOPT1ET1TTPVQDA1VRSQRGRV1G  
 RRG1Y1P1PGRSP1MDSCV1S1CEY1DAGCA1NED1TAETS1VRL1AR1L1PGLVQVQ  
 D1HLE1FVS1TGL1H1DAHF1QS1QOAGND1PY1Y1QAV1CARAQAP1PSSVDQWKK  
 L1RL1KPT1LHG1T1YLR1G1VQNEV1L1TPI1Y1MAC1NSAD1LEV1TSMV1LVGVLL  
 ALQ1K1TGS1V1TGR11LSGR1PAV1DRE1V1R1FDBMS1CASH1P1Y1EOM1OLAE  
 OYKQ1K1LGL1Q1TASQ1KABAA1PVS1WKE1WAL1EAK1WMMN1F1SG1QV1LAG1SL1P  
 PA1AS1MA1FAS1T1SPL1T1QST1L1P1L1G1W1AAQ1A1PP1SASA1P1GAG1AGNAV1S1  
 GLK1V1D1LAG1YAG1VAG1VAFK1VMS1GEMP1STED1VNL1P1L1SP1GAL1VGVW1GA  
 L1RRH1V1G1EG1V1WMMNR11APAS1RGNH1SV1THY1PVS1DAAV1R1T1LSS1T1VOLL  
 RLH1W1NED1C1P1CS1G1S1W1R1D1V1L1C1V1D1P1K1C1S1N1T1H1G1F1P1VNA1V1T1G1C1P1S  
 GW1RG1D1M1H1T1C1P1CGA1Q1TGH1K1VNG1SM1R1V1G1P1R1T1S1N1T1H1G1F1P1VNA1V1T1G1C1P1S  
 AP1NY1SK1AL1W1V1AE1VE1Y1V1T1G1D1P1HY1T1G1M1T1DN1K1C1P1Q1P1AE1P1F1E1V1D1G1V1L1R  
 YAP1CK1L1RE1V1T1Q1V1GL1N1Y1V1GS1Q1P1CB1P1D1V1AL1T1S1ML1D1PSH11TAET1AK1RL  
 YAP1CK1L1S1S1S1S1S1S1S1S1K1AT1T1R1H1D1SP1AD11E1AN1L1R1M1O1MG1N11TR1V1SE1N  
 K1V1LD1S1F1E1P1L1AE1DER1E1V1S1AE1R1K1SR1F1PAL1P1W1AR1P1N1P1L1S1S1W1Q1D1Y  
 I1P1V1D1S1P1G1P1P1T1KAP11P1P1RR1K1TV1L1T1S1V1S1AL1E1AT1K1FG1S1S1E1S1AV1D1S  
 AT1AP1D1Q1F1D1AG1S1D1V1S1S1W1P1L1E1G1R1D1P1L1D1S1G1S1W1T1S1E1B1E1S1D1V1V1C1S1M  
 Y1T1W1T1G1L1T1P1CA1B1E1S1K1P11N1AL1NS1L1RH1N1V1Y1AT1T1R1S1S1Q1B1K1Q1F1D1R1Q1L1  
 D1HY1K1E1L1K1AK1S1T1K1AK1L1S1VE1E1K11T1P1H1S1R1K1S1FG1YAG1D1V1N1L1S1G1V1H1  
 R1SV1T1D1L1ED1T1D1T11INAK1N1E1V1C1V1Q1B1K1G1R1K1P1AR11V1P1D1Q1V1R1C1K1MAL1D  
 V1W1T1L1P1Q1A1W1S1T1P1G1T1S1Q1R1VE1T1P1L1N1W1K1S1R1P1D1T1F1C1D1S1T1E1V1D1L1RT  
 B1S1Y1Q1C1D1LA1P1AR1Q11RS1L1ER1Y1V1G1P1L1N1S1K1G1N1C1Q1Y1R1C1R1AS1G1V1L1T1S1C1G1N1L  
 T1Y1L1K1AS1C1BA1K1Q1D1C1V1L1V1C1G1D1V11Y11CES1AQ1D1E1DA1S1L1RV1E1P1E1AT1R1S1V1S1P1G  
 DP1Q1P1Y1D1E1L1T1S1C1S1N1S1VA1H1D1AS1K1R1V11T1D1R1D1T1P1L1ARA1W1E1T1AR1S1T1P1N1S1L  
 G1N1I1M1Y1AT1W1ARM11W1F1S1C1R11Y11L1A1Q1B1E1K1AL1D1C1Q1Y1G1C1S1E1P1D1Q111OR1H1L  
 G1S1AS1L1H1S1Y1S1G1B1N1R1AS1K1L1G1V1P1R1V1RL1R1AS1V1R1K1L1S1Q1G1R1A1T1Q1K1Y1L  
 FN1W1AK1V1K1L1T1P11P1A1S1O1L1D1S1N1W1F1V1AG1S1G1D1Y1H1S1L1S1R1P1R1W1F1M1L1C1L1L1S1V1  
 G1Y1L1P1R1"

	Query Match	90.7%	Score 438.2	DB 14	Length 9379	
	Best Local Similarity	95.2%	Prod. No. 1.4e-101			
	Matches 452	Conservative	0	Mismatches 23	Indels 0	Gaps 0
y	2	TGCCGGGTGCTCTTCTCTATCTTCCTTCTGGCCCTGCTGCTGCTGCTGACCATCCAG	61			
b	835	TGCCGGGTGCTCTTCTCTATCTTCCTTCTGGCTGCTGCTGCTGCTGACCATCCAG	894			
y	62	CTTTCGGCTTATGAAGTGGCGCAACGTGTCGGGGGTGTACCATGTGCACGAACGACTGCTCCA	121			
b	895	CTTTCGGCTTATGAAGTGGCGCAACGGTCCGGGGGTGTACCATGTGCACGAACGACTGCTCCA	954			
y	122	ACTCAAGCATGTGTATGAGGCAGCGGACATGATCATGCACACCCCGGGTGGTGCCCT	181			
	955	ACTCAAGCATTTGTGTATGAGGCAGCGGACATGATCATGCACACCCCGGGTGGTGCCCT	1014			









Db 1207 AGCTGTTACCTTCTCGCCTCGCGCATGAGACAGTGCAGGACTGCCTCAATCT 1266  
Qy 422 ATCCCGGCCATGTATCAGGTCAACGCGATGGCTTGGGATATGATGATGACTGGTC 476  
Db 1267 ATCCCGGCCATTTATCAGGTCAACGCGATGGCTTGGGATATGATGATGACTGGTC 1321

Search completed: June 16, 2004, 11:39:17  
Job time : 1474.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 181.729 Seconds  
(without alignments)  
11290.892 Million cell updates/sec

Title: US-09-899-303a-9

Perfect score: 483

Sequence: 1 ATGCCGGTGTCTTCTTC.....TGATGAATGCTGCTTAATAG 483

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	483	100.0	483	2 AAT12707	Aat12707 HCV E1 CO
2	483	100.0	483	6 AAL48916	Aal48916 Hepatitis
3	483	100.0	483	9 ADD5517	Add5517 Hepatitis
4	474.2	98.2	480	2 AAT12708	Aat12708 HCV E1 CO
5	474.2	98.2	480	6 AAL48917	Aal48917 Hepatitis
6	474.2	98.2	480	9 ADD5519	Add5519 Hepatitis
7	435	90.1	1880	2 AaQ24467	AaQ24467 NANB hepa
8	435	90.1	2187	2 ABA03491	Aba03491 Cuticle p
9	435	90.1	2540	2 AaQ33889	AaQ33889 NANB hepa
10	435	90.1	2540	2 AaQ63753	AaQ63753 NANBHV ge
11	433.4	89.7	1863	2 AaQ15363	AaQ15363 Fragment
12	431.8	89.4	2540	2 AaQ29628	AaQ29628 Hepatitis
13	430.2	89.1	9595	2 AaX24843	Aax24843 Infection
14	430.2	89.1	9595	4 AaF23492	Aaf23492 Infection
15	430.2	89.1	9595	4 AaC86939	Aac86939 Nucleotid
16	430.2	89.1	9599	2 AaX24833	Aax24833 Infection
17	428.6	88.7	577	2 AaQ35081	AaQ35081 HCV envel
18	428.6	88.7	2187	2 ABA03492	Aba03492 Cuticle p
19	425.8	88.2	580	2 AaQ11076	AaQ11076 Fragment
20	425.8	88.2	580	2 AaZ07647	Aaz07647 HCV J1 B
21	425.4	88.1	577	2 AaQ35085	AaQ35085 HCV envel
22	425.4	88.1	580	2 AaQ79750	AaQ79750 Hepatitis
23	425.4	88.1	580	2 AaX00401	Aax00401 Hepatitis

#### ALIGNMENTS

##### RESULT 1

AAT12707

ID AAT12707 standard; DNA; 483 BP.

XX AAT12707;

AC AAT12707;

DT 23-SEP-1996 (first entry)

DE HCV E1 construct HCC112A.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX ss.

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, De Martynoff G, Buysse M;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins

XX - in presence of di: sulphide bond cleavage agent, to produce proteins

XX suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1

XX and E2 protein coding sequence constructs. These sequences are included

XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.

XX The recombinant proteins can then be isolated using a method of the

XX invention. In the method, the envelope proteins are purified by carrying

XX out a disulphide bond cleavage, or a reduction step with a disulphide

XX bond cleavage agent, after lysis of recombinant host cells. The

XX constructs containing the purified HCV envelope proteins can be used for

XX vaccinating humans against HCV, for in vitro detection of HCV antibodies

XX in a sample, and in a serotyping assay for detecting one or more

Aax26733 Consensus  
Aax26728 Consensus  
Aax79772 Hepatitis  
Aax26739 Consensus  
Aax60672 Fragment  
Aal55222 Plasmid p  
Aav60673 Fragment  
Aaq64069 Non-A, no  
Aat30387 5'UTR/COR  
Aaq64068 Non-A, no  
Aat30386 5'UTR/COR  
AaQ27160 NANB hepa  
Aat12704 HCV E1 CO  
Aal48913 Hepatitis  
Aad55634 Hepatitis  
Aav60668 Fragment  
Aaq35086 HCV envel  
Aaq79758 Hepatitis  
Aaq79760 Hepatitis  
Aaq79759 Hepatitis  
Aat12705 HCV E1 CO  
Aal48914 Hepatitis

24 425.4 88.1 580 2 AAX26733  
25 425.4 88.1 580 2 AAX26728  
26 425.4 88.1 1249 2 AAX79772  
27 425.4 88.1 1249 2 AAX26739  
28 425.4 88.1 1562 2 AAV60672  
29 425.4 88.1 1953 7 AAL55222  
30 425.4 88.1 2829 2 AAV60673  
31 425.4 88.1 3401 2 Aaq64069  
32 425.4 88.1 3401 2 AAT30387  
33 425.4 88.1 3461 2 Aaq64068  
34 425.4 88.1 3461 2 AAT30386  
35 425.2 88.0 567 2 AaQ27160  
36 424.8 88.0 642 2 AAT12704  
37 424.8 88.0 642 6 AAL48913  
38 424.8 88.0 642 9 ADD55634  
39 424.4 87.9 1270 2 AAV60668  
40 423.8 87.7 577 2 Aaq35086  
41 423.8 87.7 580 2 Aaq79758  
42 423.8 87.7 580 2 Aaq79760  
43 423.8 87.7 580 2 Aaq79759  
44 423.8 87.7 795 2 AAT12705  
45 423.8 87.7 795 6 AAL48914



[illegible]

D	B		361	CAGCTGTTACCTTCTCACCCTCGCCGGCATCAACAGATACAGGACTGCACCTGCTCAATC	480
Q	Y		421	TATCCCCGCCAATGATATCAGGTACACCGCATGCTTTGGGATATGATGATGAACATGGTCCTAA	480
D	B		421	TATCCCCGCCAATGATATCAGGTACACCGCATGCTTTGGGATATGATGATGAACATGGTCCTAA	480
Q	Y		481	TAG 483	
D	B		481	TAG 483	
 RESULT 4 AAT12708					
I	D	AAT12708 standard; DNA; 480 BP.			
X	X	AC	AAT12708;		
X	X	AC			
X	X	23-SEP-1996 (first entry)			
X	X	HCV E1 construct HCC113A.			
D	E	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.			
X	X	Hepatitis C virus.			
O	S	WO9604385-A2.			
P	N	15-PER-1996.			
P	D	31-JUL-1995; 95WO-EP003031.			
P	P	29-JUL-1994; 94EP-00870132.			
F	R	(INNO-) INNOGENETICS NV.			
X	X	Maertens G, Bosman F, De Martynoff G, Buyse M;			
X	X	WPI; 1996-129401/13.			
D	R	XX			
P	T	Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins			
P	T	- in presence of di:sulphide bond cleavage agent, to produce proteins			
P	T	suitable for direct use in vaccines or diagnostic assays of HCV.			
P	T	Claim 23; Fig 21; 146pp; English.			
X	S	AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1			
X	X	and E2 protein coding sequence constructs. These sequences are included			
C	C	in vectors for the production of recombinant E1, E2, and E1/E2 proteins.			
C	C	The recombinant proteins can then be isolated using a method of the			
C	C	invention. In the method, the envelope proteins are purified by carrying			
C	C	out a disulphide bond cleavage, or a reduction step with a disulphide			
C	C	bond cleavage agent, after lysis of recombinant host cells. The			
C	C	constructs containing the purified HCV envelope proteins can be used for			
C	C	vaccinating humans against HCV, for in vitro detection of HCV antibodies			
C	C	in a sample, and in a serotyping assay for detecting one or more			
C	C	serological types of HCV present in a biological sample. The constructs			
C	C	can also be immobilised on a solid substrate and incorporated into a			
C	C	reversed phase hybridisation assay for determining the presence or the			
C	C	genotype of HCV. The new purification method preserves the conformation			
C	C	of the recombinantly expressed E1, E2 and E1/E2, and eliminates			
C	C	contaminating proteins. Antigens isolated using this method are more			
C	C	reactive with human sera than those isolated by known techniques			
X	X	Sequence 480 BP; 85 A; 149 C; 123 G; 123 T; 0 U; 0 Other;			
S	Q	Query Match 98.2%; Score 474.2; DB 2; Length 480;			
		Best Local Similarity 99.4%; Pred. No. 1 se-128;			
		Matches 47%; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Q	Y	1 ATGCCCGGTGCTCTTCTCTATCTTCCTCTTGCCCTGCTGCTCTGTCGACCATCCA	60		

















GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 1298.46 Seconds  
(without alignments)  
11108.065 Million cell updates/sec

Title: US-09-899-303a-9  
Perfect score: 483  
Sequence: 1 ATGCCCGTGTCTTCTTC.....TGATGAACGTGCTCTTAATAG 483

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.2	13.9	488	9	AV755731
C 2	55.6	11.5	492	9	AV758366
C 3	43.2	8.9	534	14	CF846043
C 4	41.6	8.6	664	28	BZ645446

C 5	41.6	8.6	741	29	CG300832	CG300832
C 6	41.6	8.6	794	29	CG213867	CG213867
C 7	41.6	8.6	995	29	CG34840	CG34840
C 8	40.8	8.4	761	29	CG373319	CG373319
C 9	40.6	8.4	526	9	AL825643	AL825643
C 10	40.2	8.3	399	9	AV638521	AV638521
C 11	40.2	8.3	434	9	AV637507	AV637507
C 12	40.2	8.3	440	9	AV637983	AV637983
C 13	40.2	8.3	450	9	AV637259	AV637259
C 14	40.2	8.3	451	9	AV637328	AV637328
C 15	40.2	8.3	451	9	AV637643	AV637643
C 16	40.2	8.3	453	9	AV634724	AV634724
C 17	40.2	8.3	454	9	AV637050	AV637050
C 18	40.2	8.3	456	9	AV635382	AV635382
C 19	40.2	8.3	473	9	AV632765	AV632765
C 20	40.2	8.3	481	9	AV635503	AV635503
C 21	40.2	8.3	485	9	AV632811	AV632811
C 22	40.2	8.3	506	9	AV392445	AV392445
C 23	40.2	8.3	508	9	AV634095	AV634095
C 24	40.2	8.3	526	9	AV641895	AV641895
C 25	40.2	8.3	533	9	AV638125	AV638125
C 26	40.2	8.3	537	9	AV632335	AV632335
C 27	40.2	8.3	588	9	AV387329	AV387329
C 28	40.2	8.3	983	29	CG34833	CG34833
C 29	40.2	8.3	1186	13	EX421743	EX421743
C 30	39.4	8.2	1201	9	AL565958	AL565958
C 31	39.2	8.1	624	14	CD206870	CD206870
C 32	39.2	8.1	656	14	CB924688	CB924688
C 33	39.2	8.1	497	9	AV633658	AV633658
C 34	39.2	8.1	610	14	CB657655	CB657655
C 35	39.2	8.1	856	28	BZ578381	BZ578381
C 36	39.2	8.1	872	28	BZ555011	BZ555011
C 37	38.4	8.0	645	29	CNS01213	CNS01213
C 38	38.4	8.0	734	29	CC916826	CC916826
C 39	38.4	8.0	742	29	CC734434	CC734434
C 40	38.4	7.9	771	28	BZ530934	BZ530934
C 41	38.4	7.9	838	29	CC734426	CC734426
C 42	37.8	7.8	490	9	AV634529	AV634529
C 43	37.8	7.8	705	14	CA618797	CA618797
C 44	37.8	7.8	1039	13	EX415186	EX415186
C 45	37.6	7.8	309	12	BI098866	BI098866

## ALIGNMENTS

RESULT 1  
AV755731/c  
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000  
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.  
ACCESSION AV755731  
VERSION AV755731.1 GI:10913579  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukarya; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 488)  
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
Qin, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,  
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and  
Chen, Z.  
Homo sapiens cDNA BM clones  
Unpublished (2000)  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source

Location/Qualifiers  
1. .488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BMPAK303"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/clone\_lib="BM"  
/note="Vector: pTriplEx2; Site\_1: sfIIA; Site\_2: sfIIB"

## ORIGIN

Query Match 13.9%; Score 67.2; DB 9; Length 488;  
Best Local Similarity 67.0%; Pred. No. 2.8e-06;  
Matches 126; Conservative 0; Mismatches 58; Indels 4; Gaps 2;

QY 292 GTGGGGCTGCTTCTGCTTCCTGCTATGATGAGGGGATCTCTGGGATCTGTTTC 351  
DB 472 GTGGGTGTCACACTCGCTGCTCAGCTCTCTAGTGTGGGACCTCTGCGACGGAGTGATG 413  
QY 352 CTGTTTCCAGCTGTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAAC 411  
DB 412 CTTCAGTTCAGTATCA---TCTGGCTCAGCACCATGAGTTGTGATGATGCAAC 356  
QY 412 TGCTCAATCTATCCCGCCCATGATCAGGTACCGCATG-GCTTGGGATATGATGATGAA 470  
DB 355 TGCTCCATCTATCTCTGGGCCCATCACTGGACACCGTATGAGCATGGGACATGATGAA 296  
QY 471 CTGGTCTCT 478  
DB 295 CTGGTCTG 288

RESULT 2  
AV758366/C

LOCUS AV758366 BM Homo sapiens cDNA clone BMFAKA03 5', mRNA linear EST 19-OCT-2000  
DEFINITION AV758366  
ACCESSION AV758366  
VERSION AV758366.1 GI:10916214  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 492)  
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
Gu, Y., Li, N., Qian, B., Liu, P., Qu, J., Gao, X., Cheng, Z., Xu, Z.,  
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and  
Chen, Z.

TITLE  
JOURNAL

COMMENT Homo sapiens cDNA BM clones  
Unpublished (2000)  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922  
Email: hanzg@hgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source

Location/Qualifiers  
1. .492  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BMFAKA03"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/clone\_lib="BM"  
/note="Vector: pTriplEx2; Site\_1: sfIIA; Site\_2: sfIIB"

## ORIGIN

Query Match 11.5%; Score 55.6; DB 9; Length 492;  
Best Local Similarity 65.9%; Pred. No. 0.0024;  
Matches 112; Conservative 0; Mismatches 54; Indels 4; Gaps 2;  
QY 308 TCTGTTCCGCTATGTACGTTGGGGGATCTCTGGGATCTGTTTCTTCTTCCAGCTGT 367  
DB 457 TGTGATCAGCTACTACGTTGTGGACCTCTGCGTTGGGATATCGTTGAGCCCACTGA 398  
QY 368 TCACCTTCTCAGCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCTATCCCG 427  
DB 397 TTA---TCTCTCAGCAGCAACATGTTTGTGCAAGAATGCAACTGCTATTCTATCTG 341  
QY 428 GCATGTATCAGGTC-ACCGCATGCTTGGGATATGATGATGATGATGCTGGTC 476  
DB 340 GCTGCATCATCGCTACAGTATGCGATAGGCTATGATGATGATGATGATGCTGTC 291

RESULT 3  
CF846043

LOCUS CF846043  
DEFINITION CF846043 534 bp mRNA linear EST 30-OCT-2003  
ACCESSION CF846043  
VERSION CF846043.1 GI:38061697  
KEYWORDS EST.  
SOURCE Phytophthora sojae  
ORGANISM Phytophthora sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
Phytophthora.  
Tyler, B.  
Tyler, B. Not Published  
Unpublished (2003)  
Contact: Tyler B  
Tyler lab

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

FEATURES  
source

1. .534  
/organism="Phytophthora sojae"  
/mol\_type="mRNA"  
/db\_xref="taxon:67593"  
/clone="eHB036B09"  
/tissue\_type="mycelium"  
/cell\_line="P6497"  
/dev\_stage="48 hr. post infection stage"  
/lab\_host="Soybean plant"  
/clone\_lib="USDA-IPAFS:Expression of Phytophthora sojae  
genes during infection and propagation\_sHB"  
/note="Vector: pBK-CMV, Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 8.9%; Score 43.2; DB 14; Length 534;  
Best Local Similarity 47.4%; Pred. No. 3.4;  
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;  
QY 70 TATGAAGTGGCAACGTTTCGGGGGTGTACCATGTACAGACGACTGCTCACTCAAGC 129  
DB 200 TACGGCGTGGCGGAGATTACCGGTATCCGATGGGCTTCGCGGCTTCTACAACTGGACC 259  
QY 130 ATAGTGTATGAGGCGAGCATGATGATGCACACCCCGGGTGGCGCTTGGTTGG 189  
DB 260 TCGATGGACCAAGAGAGGAGGCCCGCCCATGCTGACCCCGGAGGTTGGCCACATTCAC 319



190	GAGGGCAACTCTCTCCGTTGCTGGGTGCGGCTACTCCACGCTCGCGCGCAGGAACGCC	249
Qy		
320	CACTACGGCGGACCACTCTCGGCTCGAACCGTGGTGGCTTCGACGTGGACAGAATTATC	379
Db		
250	AGCTTCCCCACAAACGACAATACGACGCCACGTCGATTGTGCTGGGGGCTGCTGCTTTTC	309
Qy		
380	AAC TTCTCGACGACGACGCGCGCTCTCGCAGGTATACGTGATCGCGCGTGAACGCCACCCAC	439
Db		
310	TGTTCCGCTATGTACGTGGGGGATCTCTGGCG	341
Qy		
440	CGTGC CGCCAAAGATCTCGGAGGAGTGC	471
Db		

RESULT 4	
B2645446/c	
LOCUS	
DEFINITION	B2645446 664 bp DNA linear GSS 29-JAN-2003
	OC6CBJ66TC ZM_0.7_1_5_KB Zea mays genomic clone ZMWBMa013P04,
	genomic survey sequence.
ACCESSION	B2645446
VERSION	B2645446.1 GI:28107610
KEYWORDS	GSS.

ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 664)
TITLE	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics

**TIGR**  
**9712 Medical Center Drive, Rockville, MD 20850, USA**  
**Contact: Cathy Whitelaw**  
**Tel: 301-838-5843**  
**Fax: 301-838-0208**  
**Email: whitelaw@tigr.org**

**FEATURES**  
**SOURCE**

## ORIGIN

Query Match	8.6%	Score 41.6;	DB 28;	Length 664;
Best Local Similarity	50.5%	Pred. NO. 9.3;		
Matches 101;	Conservative	0;	Mismatches 99;	Indels 0;
Gaps 0;				

Qy	179	CCTGCTTGGGAGGGCAATCTCTCCGCTTGCTGGTGGCGGTCACTCCACGCTCGCGG	238
Db	206	CGTACTCGCGTGGGAGCGCTCCACGAAGTGGTAGAAGCCCTCGCGCTGTTTGGCGG	147
Qy	239	CCAGGAAGCCGACGCTGCCCAACAGCAATACGACCGCCATGATTGCTCGTTGGGG	298
Db	146	ACAGGAAGCCGACGCTGCCCAAGCGACCAACACGACGGGTACCCCTTCTCGGTGTGG	87

Qy 299 CTGCTGCTTTCTGTTCCGCT 318  
||| | |||  
pb 86 GAGGTGGTGGTGGTGGCT 67

RESULT 5  
CG300832/c

CG300832	741 bp	DNA	linear	GSS 25-AUG-2003
OGWJO73TV	0.7 1.5 KB	Zea mays	genomic clone	ZMMBMa0603M01,
DEFINITION	genomic survey sequence.			

ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Professionalism	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

FEATURES  
source

## ORIGIN

Query Match 8.6%; Score 41.6; DB 29; Length 741;  
Best Local Similarity 50.5%; Pred. No. 9.7;

**Qy** 179 CCTCGTTGCGGAGGCACCTCCTCCGTTCGGTGGCGCTCACTCCCACGCTCCGCG 238  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**b4** 202 GCTATGCTGCGCGGCG 164  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
203 GCTATGCTGCGCGGCG 164  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 239 CCAGGAAGCCAGCGTCCCAACAGACAATACGACGCCACGTGCGATTTCCTCGTTGGGG 298  
pb 143 ACAGGABGACGAGGCTCTCCACGAGGCTACCCCTTCCTCGTTGGGTG 104

Qy 299 CTGCTGCTTCTGTCCGCT 318  
||| |  
nb 103 GTGTCGTGGTGTGCT 84

RESULT 6  
CG213867/C

CcZr13867/C
LOCUS
CG213867
OVCGRS3TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0506I10,
DEFINITION linear DNA 794 bp
Zea mays genome survey sequence.

**Zeal mayas**  
Bukarvota: Viridiplantae: Streptophyta: Embryophyta; Tracheophyta; Zeal mayas

**Zeal mayas**  
Bukarvota: Viridiplantae: Streptophyta: Embryophyta; Tracheophyta; Zeal mayas

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 794)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR53TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..794

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0506110"

/clone\_lib="ZM\_0.7\_1.5\_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

#### ORIGIN

Query Match 8.6%; Score 41.6; DB 29; Length 794;  
Best Local Similarity 50.5%; Pred. No. 10;  
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGCGAGCATATCATGCACACCCCGGTCGTC 178  
DB 397 CGACGAGACGACAGTGGTGGCCCAAGAGCGAGTGGAGCGGGGCCGGAC 338  
QY 179 CTTGGTTCGGGAGGCAACTCTCCCGTGTCTGGGTGCGCTCACTCCACGCTCGCG 238  
DB 337 CGTACTCGGTGGGAGCCTCCAGAGTGTAGAGCCCTGCGCTGTGTTGGCG 278  
QY 239 CCAGAACGCGAGCTCCCAAGACATAGAGCGCACGTCGATTTGCTGTTGGG 298  
DB 277 ACAGGAGCGCCAGCGCTCTCCAGAGCGACCAACAGCGGCTACCCCTTCTCGGTG 218  
QY 299 CTGCTGCTTCTGTTCGCT 318  
DB 217 GTGGTGGTGGTGGTGTCT 198

RESULT 7  
CC634840/C  
LOCUS  
DEFINITION  
CC634840TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBma050608,  
genomic survey sequence.

ACCESSION  
CC634840

VERSION  
CC634840.1 GI:32012237

KEYWORDS  
GSS.

SOURCE  
Zea mays

ORGANISM  
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 995)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR40TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..995

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma050608"

/clone\_lib="ZM\_0.7\_1.5\_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

#### ORIGIN

Query Match 8.6%; Score 41.6; DB 29; Length 995;  
Best Local Similarity 50.5%; Pred. No. 11;  
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGCGAGCATATCATGCACACCCCGGTCGTC 178  
DB 864 CGACGAGACGACAGTGGTGGCCCAAGAGCGAGTGGAGCGGGGCCGGAC 805  
QY 179 CTTGGTTCGGGAGGCAACTCTCCCGTGTCTGGGTGCGCTCACTCCACGCTCGCG 238  
DB 804 CGTACTCGGTGGGAGCCTCCAGAGTGTAGAGCCCTGCGCTGTGTTGGCG 745  
QY 239 CCAGAACGCGAGCTCCCAAGACATAGAGCGCACGTCGATTTGCTGTTGGG 298  
DB 744 ACAGGAGCGCCAGCGCTCTCCAGAGCGACCAACAGCGGCTACCCCTTCTCGGTG 685  
QY 299 CTGCTGCTTCTGTTCGCT 318  
DB 684 GTGGTGGTGGTGGTGTCT 565

RESULT 8  
CG373319/C  
LOCUS  
DEFINITION  
CG373319TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBma0723H15,  
genomic survey sequence.

ACCESSION  
CG373319

VERSION  
CG373319.1 GI:34290586

KEYWORDS  
GSS.

SOURCE  
Zea mays

ORGANISM  
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

1 (bases 1 to 761)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: CG1B144TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..761

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0723H15"

#### FEATURES

source

```

/clone lib="ZM_0.7-1.5_KB"
/note="vector: pBCKS-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      8.4%; Score 40.8; DB 29; Length 761;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 129; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 41 TGTCTGTCTGACATACAGCTTCCTCGCTTATGAAGTGCAGCAACGTTCCGGGCTGTACC 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 423 TGTCCGGGGCACGGCGCCGCCGCTGTGGCTGTGTGCTTACTTCTTCTGTTGGCG 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 101 ATGTACAGAACGACTCTCCAATCAAGCATAGTGTATGAGGAGGAGGACATGATCATGC 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 363 CCAACTCGCAGTGTTCGCGCAACACCGCGCGCTGTGTGACCTCGTCAAGAACTTCCTCCG 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 161 ACACCCCGGCTGCTGCCCTCGTGGGAGGCACTCTCTCCGTTGCTGGTGGCGC 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 303 ACAGCCCGGGGTGTGTCTGGCATCTCAAGGGCTTCGTGCGCCTCAGCGCGCGCTGT 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 221 TCACTCCACGCTCGCGGCGAGAAACGAGCTGCCCAACAGCAATACGACGCCACG 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 243 ACAGGAGCTGTACTGTGGCTGTACGGGGGAGAGCGAGCGAGTGTCTGATCTCTCTCG 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 281 TCGATTGCTGCTGGGGTGTCTGCTTTCTGTTCG 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 TCGCTCGCTCCCGCGCGGTGTCTCGTGTCTG 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AL825643/c
LOCUS      AL825643 p:234 Triticum aestivum cDNA clone A09_p234_plate_14, mRNA
DEFINITION
ACCESSION  AL825643
VERSION     AL825643.1 GI:21837164
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 526)
AUTHORS     Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
            Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
            Edwards,K.
TITLE       A BBSRC-funded wheat EST resource for the academic community
JOURNAL     Unpublished (2002)
COMMENT     Contact: Barker G
            Institute of Arable Crop Research
            Long Ashton, Bristol BS41 9AP United Kingdom.
FEATURES    Location/Qualifiers
             1..526
                /organism="Triticum aestivum"
                /mol_type="mRNA"
                /cultivar="mercia"
                /db_xref="taxon:4565"
                /clone="A09_p234_plate_14"
                /tissue_type="drought stressed seedlings"
                /dev_stage="21 days old"
                /clone_lib="p:234"

ORIGIN
Query Match      8.4%; Score 40.6; DB 9; Length 526;
Best Local Similarity 57.5%; Pred. No. 15;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 150 CATGATCATGCACACCCCGGCTGCTCGCTCGCTTCCGGAGGGCACTCTCTCCGTTG 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 151 CTTCTCTCCGAAACCCCGGGGTGATCAGTGTGATTTCTTCGGCAGCCCTTCTTGCC 92
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 210 CTGGGTGGCGCTCACTCTCCACGCTCGCGGCGAGGACGCGCTCCCAACAGCAAT 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 CTGGCGGGCGCGCGCGCACGCGGTTCGGCGCACAGGAACACCGACCGCGCGCGCGAC 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
270 ACACGCG 276
    |||
31 CACACCC 25

RESULT 10
AV638521
LOCUS      AV638521 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION
ACCESSION  AV638521.1 GI:10781841
VERSION     AV638521
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
ORGANISM    Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadales; Chlamydomonas.
REFERENCE   1 (bases 1 to 399)
AUTHORS     Asamizu,Y., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
TITLE       Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL     DNA Res. 7 (5), 305-307 (2000)
MEDLINE     20539644
PubMed      11089912
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES    Location/Qualifiers
             1..399
                /organism="Chlamydomonas reinhardtii"
                /mol_type="mRNA"
                /strain="C9"
                /db_xref="taxon:3055"
                /clone="HC087d07_r"
                /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                XhoI; The cDNA library was constructed from cells cultured
                in a medium with bubbling air containing 5% carbon
                dioxide"

ORIGIN
Query Match      8.3%; Score 40.2; DB 9; Length 399;
Best Local Similarity 51.4%; Pred. No. 17;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTGTCGCTGCTGACCATACCGCTTCGGCTTATGAAGTGCAGCAAGTGC 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44 CTTGAGCCCTCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 90 CGGGGTGTACCATGTTCACGAACGACTGCTCTCAACTCAAGCATAGTGTATGAGGAGCGGA 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 104 CGGGGTGTGCGCGCGGAGCAAGCTGACCAACATCACCGCGCGGAGGAGTGTGCTGCGG 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 150 CATGATCATGCACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 164 CATGGGCATCTAGCTGCTCCCGCGACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 210 C 210
DB 224 C 224

RESULT 11
AV637507
LOCUS      AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION
ACCESSION  AV637507
VERSION     AV637507
KEYWORDS    434 bp mRNA linear EST 15-DEC-2000
SOURCE      cDNA clone HC074a01_r 5', mRNA sequence.

```





Db 376 C 376

Search completed: June 16, 2004, 13:27:13  
Job time : 1300.46 secs







APPLICANT: DE MARTYNOFF, GUY  
APPLICANT: BUYSSE, MARIE-ANGE  
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 480 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..477  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..474

US-08-612-973-11

Query Match 98.2%; Score 474.2; DB 3; Length 480;  
Best Local Similarity 99.4%; Pred. No. 2.3e-125;  
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCCCGGTTGCTCTTTCTATCTTCTTGGCCCTGCTGCTGACCATACCA 60  
DB 1 ATGCCCGGTTGCTCTTTCTATCTTCTTGGCCCTGCTGCTGACCATACCA 60  
QY 61 GCTTCGCTTATGATGATGAGGCGGACATGATCATGACACCCCGGGTGGCTCC 120  
DB 61 GCTTCGCTTATGATGAGGCGGACATGATCATGACACCCCGGGTGGCTCC 120  
QY 121 AACTCAAGCATATGATGATGAGGCGGACATGATCATGACACCCCGGGTGGCTCC 180  
DB 121 AACTCAAGCATATGATGATGAGGCGGACATGATCATGACACCCCGGGTGGCTCC 180  
QY 181 TCGGTTCCGGGGGCACTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 181 TCGGTTCCGGGGGCACTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 241 AGGACCGCAGGCTCCCAACACACACACACACACACACACACACACACACACAC 300  
DB 241 AGGACCGCAGGCTCCCAACACACACACACACACACACACACACACACACACAC 300  
QY 301 GCTGCTTTCTGTTCCGCTATGATGATGAGGCGGACATGATCATGACACCCCGGGT 360  
DB 301 GCTGCTTTCTGTTCCGCTATGATGATGAGGCGGACATGATCATGACACCCCGGGT 360

QY 361 CAGCTGTTTACCTTCTCAGCTCCCGGCATCAACAGTACAGGACTGCAACTGCTCAATC 420  
DB 361 CAGCTGTTTACCTTCTCAGCTCCCGGCATCAACAGTACAGGACTGCAACTGCTCAATC 420  
QY 421 TATCCCGGCGCATGATGATGAGGCGGACATGATCATGATGATGATGATGATGATGAT 479  
DB 421 TATCCCGGCGCATGATGATGAGGCGGACATGATCATGATGATGATGATGATGATGAT 479

#### RESULT 4

US-08-927-597-11  
Sequence 11, Application US/08927597  
Patent No. 6245503  
GENERAL INFORMATION:  
APPLICANT: MAERTENS, GEERT  
APPLICANT: BOSMAN, FONS  
APPLICANT: DE MARTYNOFF, GUY  
APPLICANT: BUYSSE, MARIE-ANGE  
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
FILING DATE: 11-MAR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 480 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..477  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..474

US-08-927-597-11

Query Match 98.2%; Score 474.2; DB 3; Length 480;  
Best Local Similarity 99.4%; Pred. No. 2.3e-125;  
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCCCGGTTGCTCTTTCTATCTTCTTGGCCCTGCTGCTGACCATACCA 60  
DB 1 ATGCCCGGTTGCTCTTTCTATCTTCTTGGCCCTGCTGCTGACCATACCA 60  
QY 61 GCTTCGCTTATGATGATGAGGCGGACATGATCATGACACCCCGGGTGGCTCC 120













Db 37 TGGCCGGTGGCTTTCTCTATCTCTCTGGCTTGGCTGCTGTTGACCATCCAG 96  
QY 62 CTTCCGCTTATGAAGTGGCGCAACGTTGTCGGGGGTACCATGTACGACGACTGCTCCA 121  
Db 97 CTTCCGCTTATGAAGTGGCGCAACGTTGTCGGGGGTATACCATGTACGACGACTGCTCCA 156  
QY 122 ACTCAAGCATAGTATGAGGACGCGACATGATCATGACACCCCGGGTGGTGGCCCT 181  
Db 157 ACTCAAGCATAGTATGAGGACGCGACATGATCATGACATCTCCCGGGTGGTGGCCCT 216  
QY 182 GCGTTCCGGAGGCGCAACTCTCCGTTGCTGCTGGGTGGCGCTCACTCCACCGCTCGCGGCCA 241  
Db 217 GCGTTCCGGAGGACACAGCTCTCCGTTGCTGGGTAGGCGCTCACTCCACCGCTCGCGGCCA 276  
QY 242 GGAACGCCAGCGTCCCAACACAAATACGACGACGCGCATGCGATTTGCTGTTGGGGCTG 301  
Db 277 GGAATGCCAGCGTCCCACTACGACAAATACGACGCGCATGCGATTTGCTGTTGGGGCTG 336  
QY 302 CTGCTTTCTGTTGGCTATGATGAGTGGGGGATCTCTGGGATCTGTTGTTTCC 361  
Db 337 CTGCTTTCTGCTCCGCTATGATGAGTGGGGATCTCTGGGATCTGTTTCTGCTTCC 396  
QY 362 AGCTGTTCACTTTCTACCTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421  
Db 397 AGCTGTTCACTTTCTCGCTTCGCGCGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 456  
QY 422 ATCCCGGCCATGATCAGGTCACTGCGCGCATGAGTGGGATATGATGA 466  
Db 457 ATCCCGGCCATTTATCAGGTCACTGCGCGCATGAGTGGGATATGATGA 501

## RESULT 14

US-08-612-973-3  
; Sequence 3, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBB ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 642 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..639  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..636  
US-08-612-973-3

Query Match 88.0%; Score 424.8; DB 3; Length 642;  
Best Local Similarity 93.3%; Pred No. 2.6e-111;  
Matches 444; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1 ATGCCCGGTTGCTGCTTTCTCTATCTCTCTGTCGCCCTGCTGCTGCTGTCGACATACCA 60  
Db 1 ATGCCCGGTTGCTGCTTTCTCTATCTCTCTGCTGCTTTACTGCTGCTGCTGCAACATCCA 60  
QY 61 GCTTCCGCTTATGAAGTGGCGCAACGTTGTCGGGGGTGATCCATGTACGACGACTGCTCC 120  
Db 61 GCTTCCGCTTATGAAGTGGCGCAACGTTGTCGGGGATGTACCATGTACGACGACTGCTCC 120  
QY 121 AACTCAGCATAGTATGAGGACGCGACATGATGATGACACACCCCGGGTGGTGGCC 180  
Db 121 AACTCAAGCATTTGTTATGAGGACGCGACATGATGATGACACACCCCGGGTGGTGGCC 180  
QY 181 TGGTTCCGGAGGCGCAACTCTCTCCGTTGCTGCGGTGGCGCTCACTCCACCGCTCGCGGCC 240  
Db 181 TGGTTCCGGAGGACCACTCTTCCCGCTGCTGCGGTAGCGCTCACCCCAACGCTCGCAGCT 240  
QY 241 AGAAACGCCAGCGTCCCAACAGCAATACGACGCGCAACGCTGCGATTTGCTGTTGGGGCT 300  
Db 241 AGAAACGCCAGCGTCCCAACAGCAATACGACGCGCAACGCTGCGATTTGCTGTTGGGGCT 300  
QY 301 GCTGCTTTCTGTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 301 GCTGCTCTCTGTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 361 CAGCTGTTCACTTTCTCACTCCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC 420  
Db 361 CAGCTGTTCACTTTCTCGCTCTCGCGCATGAGACGCTGCGAGGACTGCAATGCTCAATC 420  
QY 421 TATCCCGGCCATGATCAGGTCACTGCGCATGCGCTTTGGGATATGATGATGATGATGATGAT 476  
Db 421 TATCCCGGCCACATAACAGGTCACTGCGTATGCGTTGGGATATGATGATGATGATGATGATGAT 476

## RESULT 15

US-08-927-597-3  
; Sequence 3, Application US/08927597  
; Patent No. 6245503  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBB ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,597

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..636
; US-08-927-597-3

```

```

Query Match      88.0%; Score 424.8; DB 3; Length 642;
Best Local Similarity 93.3%; Pred. No. 2.6e-111;
Matches 444; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      1  ATGCCCGGTGCTCTTTCTATCTCTCTTGGCCCTCTCTCTGTCTGACCATACCA 60
Db      1  ATGCCCGGTGCTCTTTCTATCTCTCTTGGCTTTACTTCTGTCTGTCTGACCATCCA 60

Qy     61  GCTTCCGCTTATGAAGTCGCAACGTGTCGGGGTGACCATGTCAAGAAAGACTGCTCC 120
Db     61  GCTTCCGCTTATGAGTGGCAACGTGTCGGGATGTACCATGTCAAGAAAGACTGCTCC 120

Qy    121  AACTCAGCATGTATGAGCGAGCAGCATGATGTCACACCCCGGGTGGTGCC 180
Db    121  AACTCAAGCATGTATGAGCGAGCAGCATGATGTCACACCCCGGGTGGTGCC 180

Qy    181  TGGCTTCGGGAGGCAACTCTCCGCTTGTGGGTGGGCTCACCTCCAGCTCGCGGC 240
Db    181  TGGCTTCGGGAGAACACTCTTCCGCTGTGGGTAGGCTCACCCCAAGCTCGAGCT 240

Qy    241  AGGAAGCCAGCGTCCCAACACGAATAACGACGCCACGTGATTTGCTGTTGGGCT 300
Db    241  AGGAAGCCAGCGTCCCAACACGACATAACGACGCCACGTGATTTGCTGTTGGGG 300

Qy    301  GCTGCTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTTCTGTTTCC 360
Db    301  GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTTCTGTTTCC 360

Qy    361  CAGCTGTTCACTCTCACTCGCGCATCAACAGTACAGGATCGCACTGCTCAATC 420
Db    361  CAGCTGTTCACTCTCGCTCGCGCATGAGAGCGGTGAGGATCGCAATGCTCAATC 420

Qy    421  TATCCCGGCCATGATCAAGGTCAACCGCATGGCTTGGGATGATGATGAAGTGGTC 476
Db    421  TATCCCGGCCATGATCAAGGTCAACCGTATGGCTTGGGATGATGATGAAGTGGTC 476

```

Search completed: June 16, 2004, 13:30:23  
Job time : 35.8162 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1464.4 Seconds  
(without alignments)  
14206.949 Million cell updates/sec

Title: US-09-899-303A-11  
Perfect score: 480  
Sequence: 1 ATGTCGGTTCCTCTTCCTC.....TGATGATGACTGGTAATAG 480

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ats.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_ats.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sv.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	480	100.0	480	6	A48673	A48673 Sequence 11
2	480	100.0	480	6	AR157328	AR157328 Sequence
3	480	100.0	480	6	AX452760	AX452760 Sequence
4	480	100.0	480	6	AX685012	AX685012 Sequence
5	480	100.0	480	6	AX794854	AX794854 Sequence
6	474.2	98.8	483	6	A48671	A48671 Sequence 9
7	474.2	98.8	483	6	AR157327	AR157327 Sequence
8	474.2	98.8	483	6	AX452758	AX452758 Sequence
9	474.2	98.8	483	6	AX685010	AX685010 Sequence
10	474.2	98.8	483	6	AX794852	AX794852 Sequence
11	436	90.8	9379	14	AP207766	AP207766 Hepatitis
12	432.8	90.2	1880	14	HPCSTRJ4	D00832 Hepatitis C
13	432.8	90.2	2540	6	E04260	E04260 cDNA encodi
14	432.8	90.2	2540	6	E04805	E04805 cDNA to 5'-
15	432.8	90.2	2540	6	E07391	E07391 cDNA encodi
16	432.8	90.2	9448	14	HPCJ483	D13558 Hepatitis C
17	431.2	89.8	1539	6	AR027786	AR027786 Sequence
18	431.2	89.8	1863	6	AR027783	AR027783 Sequence
19	430.8	89.8	9580	14	AF054250	AF054250 Hepatitis
20	429.6	89.5	8780	14	AF054257	AF054257 Hepatitis
21	429.6	89.5	8780	14	AF054258	AF054258 Hepatitis
22	428	89.2	8780	14	AF054255	AF054255 Hepatitis
23	428	89.2	8781	14	AF054254	AF054254 Hepatitis
24	428	89.2	9379	14	AF165052	AF165052 Hepatitis
25	428	89.2	9460	14	HPCJ491	D10750 Hepatitis C
26	428	89.2	9595	6	AR119832	AR119832 Sequence
27	428	89.2	9595	14	AF054247	AF054247 Hepatitis
28	428	89.2	9596	14	AF054249	AF054249 Hepatitis
29	428	89.2	9599	6	AR119833	AR119833 Sequence
30	426.4	88.8	577	6	E04085	E04085 gDNA encodi
31	426.4	88.8	618	14	HPCB1P12	D28929 Hepatitis C
32	426.4	88.8	8779	14	AF054251	AF054251 Hepatitis
33	426.4	88.8	8780	14	AF054252	AF054252 Hepatitis
34	426.4	88.8	9410	14	HPCJLR2	D50481 Hepatitis C
35	426.4	88.8	9595	14	AF054248	AF054248 Hepatitis
36	426.2	88.8	633	6	A48669	A48669 Sequence 7
37	426.2	88.8	633	6	AR157326	AR157326 Sequence
38	426.2	88.8	633	6	AX452756	AX452756 Sequence
39	426.2	88.8	633	6	AX685008	AX685008 Sequence
40	426.2	88.8	9377	14	AX794850	AX794850 Sequence
41	426	88.8	9435	14	AF207756	AF207756 Hepatitis
42	426	88.8	9435	14	AB049093	AB049093 Hepatitis
43	424.8	88.5	3296	14	AB008446	AB008446 Hepatitis
44	424.8	88.5	8780	14	AF054259	AF054259 Hepatitis
45	424.8	88.5	8781	14	AF054256	AF054256 Hepatitis

## ALIGNMENTS

RESULT 1	A48673	Sequence 11 from Patent WO9604385.	480 bp	DNA	linear	PAT 07-MAR-1997
A48673	LOCUS					
DEFINITION	A48673	Sequence 11 from Patent WO9604385.	480 bp	DNA	linear	PAT 07-MAR-1997
ACCESSION	A48673					
VERSION	A48673.1	GI:2302386				
KEYWORDS						
SOURCE	unidentified					
ORGANISM	unidentified					
REFERENCE	1 (bases 1 to 480)					
AUTHORS	Maertens,G., Bosman,F., De M.G. and Buyse,M.					
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE					
JOURNAL	Patent: WO 9604385-A 11 15-FEB-1996;					



[illegible]









AYCQTRGLGGLIITSLGRDNKNOVEGEVQVVSSTATQSTPLATCVANGVCVVVYHGAGSK  
 TLARGPKGPIOTWYNDVODVGLAPPPGARSITPCTCGSSDLVLYTRADVLPVRRRG  
 DSRGSLSPRVSYLVKSGGPGLLCPGSHVVGIPRAAVCTRGVAKADVPPVSEMETT  
 TRSPVETDSSPPAVPQPOVAHLEHAPTGSGSTKVPAYAAQGYKVLVLPNSVAAIL  
 GFGYMSKAGVDPNKTGRTITGAPITTSYTGKFLADGGCGAGYDINCDBPYGK  
 TDSILGIVDQAGTAGRLVAVLATAPGVSVPVPHNFBELISALNGBIPIYK  
 ALPIRAIKGGRRHLPFCHSKCKDELAALKSLGSLNVAAYTYRGLDVSVIPGDDVVVA  
 TDALMTGFTDLDVDCNTCTVTFDFSLDPTFTTTPVQDRAVSQRGRTRG  
 RRGIVFVTPGEPGSDSSVLCYDAGCAWYELTAPETSVRLAYLTPGLPVCO  
 DHELFWSVFTGLTHIDAHFLSOTKQAGDNLPLYVAYQATVCARAQAPPSWDQWKC  
 LIRLKPGLHPTLRLYLGAVONEVILTHPIIKYIMACMSADLEVTWVLVGGVLA  
 ALAYCLTGSVYVIGRILISGRPAVPDREVLYREDFEMERCASHLPITGEGMLAE  
 QPKKALGELLQASQAABAAVPEVSKRALEAFWAKMNPISIGIYLAGLSTLPGN  
 PALASIAFTASITPLTQSTLLFNLGNGVAAQDLAPPASAFVPGAGIAGAAVSI  
 GLKVLVDIAGAGVAGALVAFKMSGEMPSDELDNLPAILSPAGLVGVVCAA  
 ILRHVGEGEAGVOMNLLIAPASGNHVSPTHYPBESDAVVRVTLISLITOLLK  
 RLHWINDECTPCSGSLRDVMDWICTVLDPTKMLQSKLLPLRPLGPPFLSCQGYK  
 GWRGDSGIMHTTICPGAQITGHVKGSMRIUGPRICSNTHWGTFFPVNATGCTPSP  
 APNYSKALMRVAEEYVEVTRVGDVHYVTMTDNKCPQCPVAPPEFFTEVDGVRLR  
 YAPACKLLREBEVTFQVLNGVLYVGSQPCPEPDVAVLTSMLTDPSTHATKRLR  
 ARGSPSLASSASOLSAPSLKATCTRHDSPADLIENALLRDEMGENITRVSEEN  
 KVVILDSPEPRAREDEREVSAAEVLAKSRKFPALPTMARPDYNPPLLSEWQDPY  
 IPVVGCPPLPTKAPPPIPPRRKRTVLTSTVSALAEALATKTFGSESASVDSGV  
 ATAPDQDFDAGDSVSYSPMLSGEPDPLSDGWSVSESEASEDVCCSMS  
 YTTGALITPCAAESKLPINALSNLHRHNNVYATTSRASQKRTQFDRLQVLD  
 DHYRGVLIKBIKAKASTVKALLSVEACKLTPPHSARSKFGYKANDVNLKGAWHI  
 RSVKDLIEDTETPIDTITMAKNEVFCVQPERGGRKPARLIVFPDLGVRCVKMALYD  
 VVSTLPQAVMSSVGFQYSQGVREFLVNAWKSKECPWGSYDTRCFDSTVENDIRT  
 BESLYQCCDLAPEARQAIRSLTERLYVGGPLTNSKQNGCRASGVLYTSCGNLT  
 TCYLKASCAAPKALQDCTMLVCGDGLVVICESAGTQBDASLRVFTAMTYAPPG  
 DPPOEYDLELITSCSNVSAHDSKRVLYLTDRPTPLARAWETARSTPVNSWL  
 GNIIMYPTLWARMLITMTHFFSILLAQELKALDCQYGCYSTIEPDLQIOLRLH  
 GLSAFSLHSYSGPGEINRVASCLRGVPLRVRRLARSVRALKLSQCGRAATCGKYL  
 FNMAVTKLKLTPIPAAQSLDLNFWFVAGYSGDLYHLSLRARPWFMLCLLLLSGVV  
 GYLILPDR"

## ORIGIN

Query Match 90.8%; Score 436; DB 14; Length 9379;  
 Best Local Similarity 94.7%; Pred. No. 1.6e-101;  
 Matches 451; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 2 TGTCGGTGTCTTTCTATCTTCTCTTGGCCCTGCTGCTCTGACCAATCCAG 61  
 Db 835 TGCCGGTGTCTTTCTATCTTCTCTTGGCTCTGCTGCTGTTTGACCAATCCAG 894

Qy 62 CTTCGGCTATGAGTGGCGAACGTCGCGGTGTACATGTACAGAACGCTCTCCA 121  
 Db 895 CTTCGGCTATGAGTGGCGAACGTCGCGGTGTACATGTACAGAACGCTCTCCA 954

Qy 122 ACTCAAGCATAGTATGAGGAGCGGACATGATCATGCACACCCCGGTCGCTCC 181  
 Db 955 ACTCAAGCATAGTATGAGGAGCGGACATGATCATGCACACCCCGGTCGCTCC 1014

Qy 182 CGCTTCGGGAGCGCACTCTCCCTGCTGCGGTGGCGCTCACTCCCACTCGCGGCA 241  
 Db 1015 CGCTTCGGGAGCACTCTCCCTGCTGCGGTGGCGCTCACTCCCACTCGCGGCA 1074

Qy 242 GGAACGCCAGCGTCCCAACAGCAATACAGCGCCAGTCGATTCGCTGCGGTGG 301  
 Db 1075 GGAACGCCAGCGTCCCAACAGCAATACAGCGTCAGTCGATTCGCTGCGGTGG 1134

Qy 302 CTGCTTTCTGTCGCTATGAGTGGGGATCTCTGGGATCTCTTTCTGTTTCC 361  
 Db 1135 CTGCTTTCTGTCGCTATGAGTGGGGGACCTCTGCGGATCTGTTTCTGCTCTCC 1194

Qy 362 AGCTGTTCACTCTCTCACTCGCGGCAATCAACAGTACAGGACTGCAACTCTCAATCT 421  
 Db 1195 AGCTGTTCACTCTCTGCGCTCGCGGCAATGAGACGCTACAGGACTGCAATCTCAATCT 1254

Qy 422 ATCCCGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAATCGTAA 477  
 Db 1255 ATCCCGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAATCGTAA 1310

RESULT 12  
 HPC5TRJ4  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

HPC5TRJ4  
 Hepatitis C virus gene for structural protein, partial cds,  
 isolate:HC-J4.  
 D00832  
 D00832.1 GI:221513  
 structural protein.  
 Hepatitis C virus  
 Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.

1 (bases 1 to 1880)  
 Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T.,  
 Yoshizawa, H., Tada, F., Miyakawa, Y. and Mayumi, M.  
 The 5'-terminal sequence of the hepatitis C virus genome  
 Jpn. J. Exp. Med. 60 (3), 167-177 (1990)

2 (sites)  
 Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A.,  
 Miyakawa, Y. and Mayumi, M.  
 Nucleotide sequence of the genomic RNA of hepatitis C virus  
 isolated from a human carrier: comparison with reported isolates  
 for conserved and divergent regions  
 J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)

3 (sites)  
 Mink, M.A., Benichou, S., Madaule, P., Tiollais, P., Prince, A.M. and  
 Inchausti, G.  
 Characterization and mapping of a B-cell immunogenic domain in  
 hepatitis C virus E2 glycoprotein using a yeast peptide library  
 Virology 200 (1), 246-255 (1994)

4 (sites)  
 Hotta, H., Doi, H., Hayaishi, T., Purwanta, M., Soemarto, W.,  
 Mizokami, M., Ohba, K. and Homma, M.  
 Analysis of the core and E1 envelope region sequences of a novel  
 variant of hepatitis C virus obtained in Indonesia  
 Arch. Virol. 136 (1-2), 53-62 (1994)

54270990  
 7545932  
 These data kindly submitted in computer readable form by: Hiroaki  
 Okamoto  
 Immunology Division, Jichi Medical School  
 Kawachi-gun, Tochigi-ken 329-04  
 Japan  
 Phone: 0285-44-2111 x3334  
 Fax: 0285-44-1557.

Location/Qualifiers  
 1. 1880  
 /organism="Hepatitis C virus"  
 /mol\_type="genomic RNA"  
 /isolate="HC-J4"  
 /db\_xref="taxon:11103"  
 /note="82 bp upstream of StyI site"  
 342. .>1880  
 /note="putative"  
 /codon\_start=1  
 /product="structural protein"  
 /protein\_id="BAA00706.1"

/db\_xref="GI:221514"  
 /translation="MSTNPQKRTKNTNRPPQVKKPGGQIVGGVYLLPRRGPRLL  
 GVRATKTSERQPRQPIKARRPGRWAQPCYPMPLYGNELGWACHLLSPRG  
 SRPAGTDPFRSRNKLKVIDLTGCGPADLMGYIPLVGAPLGARALAHGVRLSD  
 GNVATNLPGCSFIFLLALSLCITIPASAYEVRNVSYIHTVNDNSSIIVYRAAD  
 VMHTPGCVPCVRDNRSCWVLTFLAARNASVPTTIRRHVDLLVGAAPFCAMV  
 VDLICVFLVSQLFTFPRHETVQDCNCSYPGHLSGHRMAMMMNSPTTALV  
 SOLLRIPQAVDMVAGAHVGLAGLYSMVGNNAKVLIVALLFAGVDETVTSGAA  
 SHVTSTLASLSPGASQRIQLVNTNGSWHINLTALNCNDSLHTFLAALFYTHRFNS



[illegible]





```
XX Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
KW liver fibrosis; ds; gene.
XX Hepatitis C virus.
OS WO2003051912-A2.
PN 26-JUN-2003.
XX 18-DEC-2002; 2002WO-EP014480.
XX 18-DEC-2001; 2001US-00020510.
PR 16-OCT-2002; 2002US-0418358P.
XX (INNO-) INNOGENETICS NV.
PA Maertens G, Depla E, Bosman F;
XX WPI; 2003-541632/51.
DR P-PSDB; ADD55520.
XX New hepatitis C virus (HCV) vaccine composition, useful for reducing
PT liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX Example 1; SEQ ID NO 11; 271pp; English.
XX The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
CC liver disease. The vaccine of the invention comprises an HCV E1 or E2
CC protein as an antigen. The HCV vaccine is useful for reducing liver
CC disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
CC present DNA sequence encodes an HCV E1 protein.
XX
SQ Sequence 480 BP; 85 A; 149 C; 123 G; 123 T; 0 U; 0 Other;

Query Match      100.0%; Score 480; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 3e-130;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCGGGTGCTCTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Db 1 ATGTCGGGTGCTCTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60

Qy 61 GCTTCGGCTTATGAAGTCGCAACGTCGCGGGGTATACCATGTCAAGAACGACTGCTCC 120
Db 61 GCTTCGGCTTATGAAGTCGCAACGTCGCGGGGTATACCATGTCAAGAACGACTGCTCC 120

Qy 121 AACTCAGCATAGTGTATGAGGCGGACGACATGATGTCACACCCCGGGTGGTGCC 180
Db 121 AACTCAGCATAGTGTATGAGGCGGACGACATGATGTCACACCCCGGGTGGTGCC 180

Qy 181 TGGCTTCGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240

Qy 241 AGGAAGCCAGCGTCCCCACACACATACGACCCACGTCGATTTGCTGTGGGGCT 300
Db 241 AGGAAGCCAGCGTCCCCACACACATACGACCCACGTCGATTTGCTGTGGGGCT 300

Qy 301 GCTGCTTCTCTCGGTATGTAAGTGGGGATCTCTCGGGATCTGTTCTGTGTTCC 360
Db 301 GCTGCTTCTCTCGGTATGTAAGTGGGGATCTCTCGGGATCTGTTCTGTGTTCC 360

Qy 361 CAGCTGTTCACTTCTACCTCGCGGACATCAACAGTACAGGACTGCAATGCTCAATC 420
Db 361 CAGCTGTTCACTTCTACCTCGCGGACATCAACAGTACAGGACTGCAATGCTCAATC 420

Qy 421 TATCCCGGCCATGATCAGTCAACGATGCTGGGATATGATGATGAATGATGATGATG 480
Db 421 TATCCCGGCCATGATCAGTCAACGATGCTGGGATATGATGATGAATGATGATGATG 480
```

RESULT 4

```
AAT12707
ID AAT12707 standard; DNA; 483 BP.
XX
AC AAT12707;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC112A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP003031.
XX
PR 29-JUL-1994; 94EP-00870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, De Martynoff G, Buysse M;
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
PT - in presence of disulphide bond cleavage agent, to produce proteins
PT suitable for direct use in vaccines or diagnostic assays of HCV.
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by carrying
CC out a disulphide bond cleavage, or a reduction step with a disulphide
CC bond cleavage agent, after lysis of recombinant host cells. The
CC constructs containing the purified HCV envelope proteins can be used for
CC vaccinating humans against HCV, for in vitro detection of HCV antibodies
CC in a sample, and in a serotyping assay for detecting one or more
CC serological types of HCV present in a biological sample. The constructs
CC can also be immobilised on a solid substrate and incorporated into a
CC reversed phase hybridisation assay for determining the presence or the
CC genotype of HCV. The new purification method preserves the conformation
CC of the recombinantly expressed E1, E2 and E1/E2, and eliminates
CC contaminating proteins. Antigens isolated using this method are more
CC reactive with human sera than those isolated by known techniques
XX
SQ Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 U; 0 Other;

Query Match      98.8%; Score 474.2; DB 2; Length 483;
Best Local Similarity 99.4%; Pred. No. 1.5e-128;
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGTCGGGTGCTCTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Db 1 ATGTCGGGTGCTCTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60

Qy 61 GCTTCGGCTTATGAAGTCGCAACGTCGCGGGGTATACCATGTCAAGAACGACTGCTCC 120
Db 61 GCTTCGGCTTATGAAGTCGCAACGTCGCGGGGTATACCATGTCAAGAACGACTGCTCC 120

Qy 121 AACTCAGCATAGTGTATGAGGCGGACGACATGATGTCACACCCCGGGTGGTGCC 180
Db 121 AACTCAGCATAGTGTATGAGGCGGACGACATGATGTCACACCCCGGGTGGTGCC 180

Qy 181 TGGCTTCGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
```











SQ Sequence 1863 BP; 333 A; 586 C; 547 G; 397 T; 0 U; 0 Other;  
 Query Match 89.8%; Score 431.2; DB 2; Length 1863;  
 Best Local Similarity 94.1%; Pred. No. 9.3e-116;  
 Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
 QY 2 TGTCGGTTGCTCTTCTCTATCTCTTGGCCCTGCTGCTGACCAATACCAG 61  
 DB 830 TGGCCGGTTGCTCTTCTCTATCTCTTGGCCCTGCTGCTGACCAATACCAG 889  
 QY 62 CTTCCCGCTTATGAAGTGGCAACGCTGTCGGGGGTACCATGTCCAGAACGACTGCTCCA 121  
 DB 890 CTTCCCGCTTATGAAGTGGCAACGCTGTCGGGGGTATACCATGTCCAGAACGACTGCTCCA 949  
 QY 122 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCTGCCCT 181  
 DB 950 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCATCTCCCGGGTGGCTGCCCT 1009  
 QY 182 GCGTTCCGGGAGGCAATCTCTCCGTTGCTGGGTGGCGCTCACTCCACAGCTCGCGGCCA 241  
 DB 1010 GCGTTCCGGGAGGCAACAGCTCCGTTGCTGGGTGGCGCTCACTCCACAGCTCGCGGCCA 1069  
 QY 242 GGAACGCCAGCGTCCCAACAGCAATACGACCGACGCTGATTTGCTGGGGCTG 301  
 DB 1070 GGAATGCCAGCGTCCCAACAGCAATACGACCGACGCTGATTTGCTGGGGCTG 1129  
 QY 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361  
 DB 1130 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1189  
 QY 362 AGCTGTTTCACTTCTCACTCGCGGCATCAAGACAGTACAGGACTGCAACTGCTCAATCT 421  
 DB 1190 AGCTGTTTCACTTCTCGCTCGCGGCATGAGACAGTACAGGACTGCAACTGCTCAATCT 1249  
 QY 422 ATCCCGGCCATGATCAGTGCACGCGATGCTTGGGATATGATGATGATGATGATGATGAT 477  
 DB 1250 ATCCCGGCCATTTATCAGGTGACCGCATGGCTTGGGATATGATGATGATGATGATGATGAT 1305

RESULT 12  
 AAQ29628  
 ID AAQ29628 standard; DNA; 2540 BP.  
 XX  
 AC AAQ29628;  
 XX  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 16-MAR-1993 (first entry)  
 XX  
 DE Hepatitis C virus HC-J4 5' region.  
 XX  
 KW Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening; PCR;  
 KW primer; polymerase chain reaction; ss.  
 XX  
 OS Hepatitis C virus.  
 XX  
 XX EP510952-A1.  
 XX  
 PD 28-OCT-1992.  
 XX  
 XX 23-APR-1992; 92EP-00303625.  
 XX  
 PR 26-APR-1991; 91JP-00191376.  
 XX  
 PA (IMMO) IMMUNO JAPAN INC.  
 XX  
 PI Okamoto H, Nakamura T;  
 XX  
 DR WPI; 1992-359137/44.  
 XX  
 XX Detection of non-A, non-B hepatitis virus - using new oligo-nucleotide  
 PT primers with nucleotide sequences corresp. to part. of the viral RNA.  
 XX  
 PS Disclosure; Page 18; 54pp; English.

XX This sequence represents the 5' region of hepatitis C virus RNA. The  
 CC original sample was obtained from human and chimpanzee plasma. RNA was  
 CC isolated from several samples and homology compared, and the respective  
 CC sequence of about 1900 - 2500 nucleotides of the 5' terminus and 1100  
 CC nucleotides of the 3' terminus determined. The 5' region (given) contains  
 CC a non-coding region of at least 340 nucleotides and a region coding for  
 CC the structural protein followed by a region coding for the non-structural  
 CC protein (none actually detailed on the sequence given in the  
 CC specification). When compared with the sequence of HCV disclosed in EP-  
 CC 388232 this sequence showed homology of 80.5%. (Updated on 25-MAR-2003 to  
 CC correct FN field.)

SQ Sequence 2540 BP; 472 A; 775 C; 741 G; 552 T; 0 U; 0 Other;  
 Query Match 89.5%; Score 429.6; DB 2; Length 2540;  
 Best Local Similarity 93.9%; Pred. No. 3.1e-115;  
 Matches 447; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
 QY 2 TGTCGGTTGCTCTTCTCTATCTCTTGGCCCTGCTGCTGACCAATACCAG 61  
 DB 847 TGGCCGGTTGCTCTTCTCTATCTCTTGGCCCTGCTGCTGACCAATACCAG 906  
 QY 62 CTTCCCGCTTATGAAGTGGCAACGCTGTCGGGGGTACCATGTCCAGAACGACTGCTCCA 121  
 DB 907 CTTCCCGCTTATGAAGTGGCAACGCTGTCGGGGGTATACCATGTCCAGAACGACTGCTCCA 966  
 QY 122 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCTGCCCT 181  
 DB 967 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCATCTCCCGGGTGGCTGCCCT 1026  
 QY 182 GCGTTCCGGGAGGCAACCTCTCTCCGTTGCTGGGTGGCGCTCACCTCCACGCTCGCGGCCA 241  
 DB 1027 GCGTTCCGGGAGGCAACAGCTCCGTTGCTGGGTGGCGCTCACCTCCACGCTCGCGGCCA 1086  
 QY 242 GGAACGCCAGCGTCCCAACAGCAATACGACCGACGCTGATTTGCTGGGGCTG 301  
 DB 1087 GGAATGCCAGCGTCCCAACAGCAATACGACCGACGCTGATTTGCTGGGGCTG 1146  
 QY 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361  
 DB 1147 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1206  
 QY 362 AGCTGTTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421  
 DB 1207 AGCTGTTTCACTTCTCGCTCGCGGCATGAGACAGTACAGGACTGCAACTGCTCAATCT 1266  
 QY 422 ATCCCGGCCATGATCAGTGCACGCGATGCTTGGGATATGATGATGATGATGATGATGAT 477  
 DB 1267 ATCCCGGCCATTTATCAGTGCACGCGATGCTTGGGATATGATGATGATGATGATGATGAT 1322

RESULT 13  
 AAX24843  
 ID AAX24843 standard; DNA; 9595 BP.  
 XX  
 AC AAX24843;  
 XX  
 DT 21-JUN-1999 (first entry)  
 XX  
 DE Infectious hepatitis C virus genotype 1b strain HC-J4 genome.  
 XX  
 KW HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;  
 KW assay; antiviral; virucide; ss.  
 XX  
 OS Hepatitis C virus.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 342..9374  
 FT /\*tag= a  
 XX  
 XX WO9904008-A2.

```
PD XX 28-JAN-1999.
PF XX 16-JUL-1998; 98WO-US014688.
PR XX 18-JUL-1997; 97US-0053062P.
PR XX 27-JAN-1998; 98US-00014416.
PR XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
PI XX WPI; 1999-132252/11.
DR XX P-PSDB; AAW98022.
DR XX
XX XX
PT PT New isolated hepatitis C virus nucleic acids - used to develop products
PT PT for the diagnosis, prevention and treatment of HCV infections and for
PT PT developing screening assays.
XX XX
PS PS Claim 3; Fig 14A-F; 126pp; English.
XX XX
CC CC The present sequence comprises the nucleic acid sequence of the genome of
CC CC infectious hepatitis C virus (HCV) genotype 1b strain HC-J4 (ATCC 209596)
CC CC that is capable of expressing this virus when transfected into cells. HC-
CC CC J4 was obtained from acute phase plasma of a chimpanzee experimentally
CC CC infected with serum containing HC-J4/91. The claimed infectious nucleic
CC CC acid sequence can be used to produce chimeric genomes (see AAX24833)
CC CC consisting of the open reading frames of infectious nucleic acid
CC CC sequences of other genotypes (including genotypes 1-6) and subtypes (such
CC CC as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also
CC CC relates to the introduction of mutations or deletions into infectious
CC CC nucleic acid sequences in order to produce an attenuated HCV virus
CC CC suitable for vaccine development. Infectious nucleic acid sequences can
CC CC also be used to produce attenuated virus via passage in vitro or in vivo
CC CC of the viruses produced by transfection of a host cell with the
CC CC infectious nucleic acid sequence. Vaccines comprising one or more
CC CC polypeptides made from the infectious nucleic acid sequence are used to
CC CC immunise mammals, especially humans, against hepatitis C. The nucleic
CC CC acid sequences can also be used to induce protective immunity against the
CC CC virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
CC CC protease) can additionally be used to develop screening assays to
CC CC identify antiviral agents for HCV
XX XX
SQ SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;

Query Match      89.2%; Score 428; DB 2; Length 9595;
Best Local Similarity 93.7%; Pred. No. 1.5e-114;
Matches 446; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 TGTCGGTGTGCTTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGTGTGACCATACAG 61
Db 847 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTGCTGCTCTGTGTGACCATACAG 906
Qy 62 CTTCCGCTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCCA 121
Db 907 CTTCCGCTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCCA 966
Qy 122 ACTCAAGCATAGTGTATGAGCGAGCAGCATGATCATGTCACACCCCGGGTGGTGCCT 181
Db 967 ACTCAAGCATAGTGTATGAGCGAGCAGCATGATCATGTCACACCCCGGGTGGTGCCT 1026
Qy 182 CGCTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGGCA 241
Db 1027 GTGTTCAGGAGGTAAACAGTCTCCGTTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCA 1086
Qy 242 GGAAAGCCAGCGTCCCAACAAACGAATACGAACGCCACGTGCAATTTGCTGTTGGGGCTG 301
Db 1087 GGAAATGCCAGGTCCTCCCACTACGAACAAATACGAACGCCACGTGCAATTTGCTGTTGGGGCTG 1146
Qy 302 CTGCTTTCTGTCGGCTATGATGCTGGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db 1147 CTGCTTTCTGTCGGCTATGATGCTGGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 1206
Qy 362 AGCTGTTTCACTTCTACCTCGCCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
```

```
Db 1207 AGCTGTTTCACTTCTGCTCGCCGATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
Qy 422 ATCCCGGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAAGTGTAA 477
Db 1267 ATCCCGGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAAGTGTAA 1322

RESULT 14
AAP23492
ID AAP23492 standard; DNA; 9595 BP.
XX
AC AAP23492;
XX
DT 21-MAR-2001 (first entry)
DE Infectious Hepatitis C virus 1b genotype.
XX
KW GBV-B; hepatitis C virus; HCV; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO2000075337-A1.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015293.
XX
PR 04-JUN-1999; 99US-0137694P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX WPI; 2001-091214/10.
XX
XX New infectious nucleic acids of the GB virus-B clone, useful for
XX indirectly studying the molecular properties of hepatitis C virus (HCV)
XX PT and in developing vaccines and therapeutics for HCV.
XX
XX Disclosure; Fig 7; 96pp; English.
XX
XX The present invention relates to GB virus-B. The nucleic acid molecules
XX of the invention are useful for indirectly studying the molecular
XX properties of hepatitis C virus (HCV). The infectious nucleic acid
XX sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
XX CC in the development of vaccines and therapeutics for HCV
XX
SQ Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 U; 0 Other;

Query Match      89.2%; Score 428; DB 4; Length 9595;
Best Local Similarity 93.7%; Pred. No. 1.5e-114;
Matches 446; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 TGTCGGTGTGCTTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGTGTGACCATACAG 61
Db 847 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTGCTGCTCTGTGTGACCATACAG 906
Qy 62 CTTCCGCTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCCA 121
Db 907 CTTCCGCTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCCA 966
Qy 122 ACTCAAGCATAGTGTATGAGCGAGCAGCATGATCATGTCACACCCCGGGTGGTGCCT 181
Db 967 ACTCAAGCATAGTGTATGAGCGAGCAGCATGATCATGTCACACCCCGGGTGGTGCCT 1026
Qy 182 CGCTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGGCA 241
Db 1027 GTGTTCAGGAGGTAAACAGTCTCCGTTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCA 1086
Qy 242 GGAAAGCCAGCGTCCCAACAAACGAATACGAACGCCACGTGCAATTTGCTGTTGGGGCTG 301
Db 1087 GGAAATGCCAGGTCCTCCCACTACGAACAAATACGAACGCCACGTGCAATTTGCTGTTGGGGCTG 1146
```





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: June 16, 2004, 09:12:53 ; Search time 1290.4 Seconds  
(without alignments)  
11108.065 Million cell updates/sec

Title: US-09-899-303A-11  
Perfect score: 480  
Sequence: 1 ATGTCGGTTCCTTTCTC.....TGATGATGACGTGTAATAG 480

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	65.8	13.7	488	9 AV755731	AV755731
C 2	54.6	11.4	492	9 AV758366	AV758366
C 3	43.2	9.0	534	14 CF846043	CF846043
C 4	41.6	8.7	564	28 BZ645446	BZ645446

C 5	41.6	8.7	741	29 CG300832	CG300832
C 6	41.6	8.7	794	29 CG213867	CG213867
C 7	41.6	8.7	995	29 CG634840	CG634840
C 8	40.8	8.5	761	29 CG373319	CG373319
C 9	40.6	8.5	526	9 AL825643	AL825643
C 10	40.2	8.4	399	9 AV638521	AV638521
C 11	40.2	8.4	434	9 AV637507	AV637507
C 12	40.2	8.4	440	9 AV637983	AV637983
C 13	40.2	8.4	450	9 AV637259	AV637259
C 14	40.2	8.4	451	9 AV637328	AV637328
C 15	40.2	8.4	451	9 AV637643	AV637643
C 16	40.2	8.4	453	9 AV634724	AV634724
C 17	40.2	8.4	454	9 AV637050	AV637050
C 18	40.2	8.4	456	9 AV635382	AV635382
C 19	40.2	8.4	473	9 AV632765	AV632765
C 20	40.2	8.4	481	9 AV635503	AV635503
C 21	40.2	8.4	485	9 AV632811	AV632811
C 22	40.2	8.4	506	9 AV392445	AV392445
C 23	40.2	8.4	508	9 AV634095	AV634095
C 24	40.2	8.4	526	9 AV641895	AV641895
C 25	40.2	8.4	533	9 AV638125	AV638125
C 26	40.2	8.4	537	9 AV632335	AV632335
C 27	40.2	8.4	588	9 AV387329	AV387329
C 28	40.2	8.4	983	29 CG634833	CG634833
C 29	40	8.3	1186	13 BX421743	BX421743
C 30	39.4	8.2	1201	9 AL565958	AL565958
C 31	39.2	8.2	624	14 CD206870	CD206870
C 32	39.2	8.2	656	14 CB924688	CB924688
C 33	39	8.1	497	9 AV633658	AV633658
C 34	39	8.1	610	14 CB657655	CB657655
C 35	39	8.1	856	28 BZ578381	BZ578381
C 36	39	8.1	872	28 BZ555011	BZ555011
C 37	38.4	8.0	645	29 CMS01213	CMS01213
C 38	38.4	8.0	734	29 CC916826	CC916826
C 39	38	7.9	742	29 CC734434	CC734434
C 40	38	7.9	771	28 BZ530934	BZ530934
C 41	38	7.9	838	29 CC734426	CC734426
C 42	37.8	7.9	490	9 AV634529	AV634529
C 43	37.8	7.9	705	14 CA618797	CA618797
C 44	37.8	7.9	1039	13 BX415186	BX415186
C 45	37.6	7.8	309	12 BI098866	BI098866

## ALIGNMENTS

RESULT 1  
AV755731/c 488 bp mRNA linear EST 19-OCT-2000  
LOCUS AV755731 BM Homo sapiens cDNA clone BMAPX03 5', mRNA sequence.  
DEFINITION AV755731  
ACCESSION AV755731  
VERSION AV755731.1 GI:10913579  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 488)  
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.  
TITLE Homo sapiens cDNA BM clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

[illegible]

Qy	190	GAGGGCAATCCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTGCGGCCGAGAACGCC	249
Db	320	CACCTACGGCGGCAACATCTCTCGGCTCGAACCGTGGTGGCTTCGACGTGCAAGAATTATC	379
Qy	250	AGCGTCCCCACAACGACAATACGACGCCACGCTCGATTTCGTGTTGGGGCTGCTGCTTTC	309
Db	380	AAC TTCCTGACGACGAACGGCGTCTCGCAGGTATACGTGATCGCGGTGACGGCACCCAC	439
Qy	310	TGTTCCCGCTATGTACGTGGGGGATCTCTCGCG	341
Db	440	CGTGCCGCCCAACAGATCTCGAGGAGTGC	471

RESULT 4	BZ645446/c
LOCUS	OQCBJ86TC_ZM_0.7_1.5_KB_Zea mays genomic clone ZMWBNa0133P04,
DEFINITION	664 bp DNA linear GSS 29-JAN-2003
ACCESSION	BZ645446
VERSION	BZ645446.1 GI:28107610
KEYWORDS	GSS.

SOURCE	ORGANISM
Zea mays	
Zea mays	
Eukaryota;	Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 664)	
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citsek, R.W., Nunberg, A., Robbins, D. and Lakey, N.	
REFERENCE	
AUTHORS	

TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other\_GSSRs: OGCBJ66TM  
Contact: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.

```

FEATURES
  source
    Class: sheared ends.
    Location/Qualifiers
      1..664
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMEM0133P04"
        /clone_lib="ZM_0.7 1.5 kb"
        /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"

```

ORIGIN	methylation induced genomic DNA library									
	Query Match	8.7%	Score 41.6;	DB 28;	Length 654;					
	Best Local Similarity	50.5%;	Pred. No. 9.3;							
	Matches 101;	Conservative 0;	Mismatches 99;	Indels 0;	Gaps 0;					
Qy	119	CGAACTCAAGCATAGTGTATGAGGACGGACATGATCATGCACACCCCGGGTGGTGC	178							
Db	266	CCGACGAGACGAACAACGTGGGTGGGCCCAAGACGAGGTGCGGAAGCAGGGGCCGGGAC	207							
Qy	179	CTTGCCCTTCGGGAGGGCAACTCTCCCTGTTGCTTGGGTGGCGCTCACTCCACACCTCGCGG	238							
Db	206	CGTACCTGCGGTGGCGAGCTCCACGAAGTGTAGAAGCCCTCGCCGCTGCTGTGGCGG	147							
Qy	239	CAGGAACGCCAGCGTCCCAACAACACAAATACGACGCCACGTGATTTGCTCGTTGGGG	298							
Db	146	ACAGGAAGGCAGCGTCTCCACGACGGACCAACACGAGGGCTACCCCTTCTCGGTGGTG	87							

Qy 299 CTGCTGCTTTCTGTTCCGCT 318  
||| | |||  
Db 86 GTGCTGCTGCTGCTGCT 67

RESULT 5	ACCESSION	ORGANISM
CG300832/c	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
----------------------	-----------------------------

FEATURES  
source

ORIGIN

Query Match	Best Local	Matches
1	Qy	
2	Db	
1	Qy	
2	Db	
2	Qy	
1	Db	
2	Qy	
1	Db	

RESULT 6	ACCESSION	ORGANISM
CG213867/c	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

CG300832 741 bp DNA linear GSS 25-AUG-2003  
 CGWJ0719V ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMWma0603M01,  
 genomic survey sequence.  
 CG300832  
 CG300832 CG300832.1 GI:34215046  
 GSS.  
 Zea mays  
 Zea mays  
 Zea mays  
 Eukaryota  
 Viridiplantae: Streptophyta; Embryophyta: Tracheophyta:  
 Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 741)  
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
 Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Contact: Cathv Whitelaw

```
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: Sheared ends.
      Location/Qualifiers
        1..741
          /organism="Zea mays"
          /mol_type="genomic DNA"
          /strain="B73"
          /db_xref="taxon:4577"
          /clone="ZMMBMA0603W01"
          /clone_lib="ZM 0.7 1.5_KB"
          /note="vector: pBGSK; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
```

Ph	8.7%;	Score	41.6;	DB	29;	Length	741;
Similarity	50.5%;	Pred. NO.	9.7;				
Conservative	0;	Mismatches	99;	Indels	0;	Gaps	0;
9	CCAACTCAAGCATAGTGTATCAGGCAGCGGACATGATCATGCACACCCCGGGTGGTGC	178					
33	CCGACGACACGACATCGTGGTGGGCCCAAGACGAGGTCCGAAGACGAGGSCCGGAC	224					
9	CCTGGTTCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTACTCCCA CGCTCGCGG	238					
33	CGTACTCGGTGGCGAGCCTCCACGAAGTGGTAGAAGCCCTGCGCGCTGCTGTTGGCGG	164					
9	CCAGGACGCCAGGTGCCCAACGACAATACGAGGCCACGTGCATTTGCTCGTTGGGG	298					
33	ACAGGAGGCCAGCGGTCTCCACAGCGACCAACACAGGGCTACCCCTTTCGGTGGTG	104					
9	CTGCTGCTTTCTGTTCCGCT	318					
	GTGCTGGTGGTGGTGGTGGT	84					

CG213867 794 bp DNA linear GSS 22-AUG-2003  
 OGVCR33TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBma0506i10,  
 genomic survey sequence.  
 CG213867  
 CG213867.1 GI:34113755 GSS.  
 Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 794)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR53TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..794

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0506110"

/clone\_lib="ZM 0.7 1.5 KB"

/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

#### ORIGIN

Query Match 8.7%; Score 41.6; DB 29; Length 794;  
Best Local Similarity 50.5%; Pred. No. 10;  
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGACGAGGACATGATCATGACACACCCCGGGTGGTC 178  
DB 397 CCAGCAGACGACACGTCGTGGTGGCGCCCAAGAGCGAGGTGCGGAGCGGGCCGGAC 338  
QY 179 CTGTGGTTCGGGAGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGG 238  
DB 337 CGTACCTCGGTGCGGAGGCTCCACGAAAGTGGTAGAGCCCTGCGCTGCTGTGGCGG 278  
QY 239 CCAGGACGCGAGGTCCTCCACACGACATGACGACGCGGAGTTCGTTGCGGG 298  
DB 277 ACAGGAAGGCGCGGCTCTCCACGACGAGCACCACCGAGCGGCTTCTCGGTGGTG 218  
QY 299 CTGCTGCTTTCTGTTCCGCT 318  
DB 217 GTGGTGGTGGTGGTGGTCT 198

#### RESULT 7

CC634840/c  
LOCUS  
DEFINITION  
CGVCB40TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0500G08,  
genomic survey sequence.

ACCESSION  
CC634840

VERSION  
CC634840.1 GI:32012237

KEYWORDS  
GSS.

SOURCE  
Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 995)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCB40TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

#### FEATURES source

Location/Qualifiers  
1..995  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBma0500G08"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

#### ORIGIN

Query Match 8.7%; Score 41.6; DB 29; Length 995;  
Best Local Similarity 50.5%; Pred. No. 11;  
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGACGAGGACATGATCATGACACCCCGGGTGGTC 178  
DB 864 CCAGCAGACGACACGTCGTGGTGGCGCCCAAGAGCGAGGTGCGGAGCGGGCCGGAC 805  
QY 179 CTGTGGTTCGGGAGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGG 238  
DB 804 CGTACCTCGGTGCGGAGGCTCCACGAAAGTGGTAGAGCCCTGCGCTGCTGTGGCGG 745  
QY 239 CCAGGAACGCGAGGTCCTCCACACGACAAATACAGCCGACGTCGATTTGCTGTTGGGG 298  
DB 744 ACAGGAAGGCGAGGCTCTCCACGACGAGCACCACCGAGCGGCTTCTCGGTGGTG 685  
QY 299 CTGCTGCTTTCTGTTCCGCT 318  
DB 684 GTGGTGGTGGTGGTGGTGGTCT 665

#### RESULT 8

CG373319/c

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

CG373319

VERSION

CG373319.1 GI:34290586

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 761)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGI44TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..761

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0723H15"

```

/clone_lib="ZK_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      8.5%; Score 40.8; DB 29; Length 761;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 129; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 41 TGTCTGTCTGTGACACACACAGCTTCGGCTTATGAAGTGCAGCACTGTCGGGGGTATCC 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 TGTCCGGGGCAGCGCGCGCGCGCTGTGGCTGTGGTGCCTTACTTCTTGTGGGCG 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 101 ATGTCAAGACAGCTGCTCAACTCAAGCAATGATGTATGAGGCGGACATGATCATGC 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 CCAACTCGCAGTGTCTCGCAACACACCGCGCGCTGTGTGACCTGCTCAAGAACTTCCCG 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 161 ACACCCCGGGTGTGGCTGTGGTGGGAGGCACTCTCCGTTGTGGTGGCGC 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 ACAGCCCGGGGTGTGGTGTGGGCACTCTCAAGGGCTTGTGGCTCAGCGGGCGGTGT 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 221 TCACTCCACAGCTCGCGGCGCAGCAAGCCAGCGTCCCAACAGCAATACAGCGCAAG 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 ACAGCAGCTGTACCTGGCCCTGTAGCGCGCAGCGAGTCTGCTGATCTCTCTCG 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 281 TCGATTGTCTGTGGGGCTGTGCTGCTTCTGTTCCG 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 TCGCTGGCTCCCGCGCGGTGTGCTGCTGTGTTCG 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AL825643/c
LOCUS
DEFINITION
AL825643 526 bp mRNA linear EST 15-JUL-2002
AL825643 p:234 Triticum aestivum cDNA clone A09_p234_plate_14, mRNA
sequence.
ACCESSION
AL825643
VERSION
AL825643.1 GI:21837164
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 526)
Edwards, D., Coghill, J., Shepherd, S., Barker, G., Parker, J., Owen, P.,
Edwards, K.
A BBSRC-funded wheat EST resource for the academic community
Unpublished (2002)
Contact: Barker G
Institute of Arable Crop Research
Long Ashton, Bristol BS41 9AP United Kingdom.
FEATURES
source
1..526
Location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="mercia"
/db_xref="taxon:4565"
/clone="A09_p234_plate_14"
/tissue_type="drought stressed seedlings"
/dev_stage="21 days old"
/clone_lib="p:234"

ORIGIN
Query Match      8.5%; Score 40.6; DB 9; Length 526;
Best Local Similarity 57.5%; Pred. No. 15;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTCGCTTGGGAGGCACTCTCTCCGGTGG 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 CCTCTCCGACACCCCGGGTGGATCAGTGGGAGTTCTTCGGCAGCCCTTCTTGGC 92
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 210 CTGGGTGGCGCTCACTCTCCACGCTCGCGGCCAGGAACGCGCTGCCCAACACGCAAT 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 CTGGGGCGCGCGCGCACCGCGGTGCGGCACAGGACACACCGACCGCGCGCGCGCAC 32
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 270 ACGAGCG 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 31 CACACCC 25
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AV638521
LOCUS
DEFINITION
AV638521 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC087d07_r 5', mRNA sequence.
ACCESSION
AV638521
VERSION
AV638521.1 GI:10781841
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE
1 (bases 1 to 399)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
20539644
PUBMED
11089912
COMMENT
Contact: Brika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..399
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC087d07_r"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"

ORIGIN
Query Match      8.4%; Score 40.2; DB 9; Length 399;
Best Local Similarity 51.4%; Pred. No. 17;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44 CTTTCACCCCTGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 90 CGGGGTGTACCATGTCACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 104 CGCGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTCGCTTGGGAGGCACTCTCTCCGGTGG 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 CATGGGCATCTAAGTCTCCCGCACCGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 210 C 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 C 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
AV637507
LOCUS
DEFINITION
AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC074a01_r 5', mRNA sequence.

```







Db 376 C 376

Search completed: June 16, 2004, 13:27:14  
Job time : 1291.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 34.6 Seconds  
(without alignments)  
7698.741 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTTCCTCTTCTC.....TGATGATGACTGGTAATAG 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgm2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgm2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgm2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgm2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgm2\_6/ptodata/2/ina/ECTUS\_COMB.seq.\*  
6: /cgm2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	3	US-08-612-973-11
2	480	100.0	480	3	US-08-927-597-11
3	474.2	98.8	483	3	US-08-612-973-9
4	474.2	98.8	483	3	US-08-927-597-9
5	431.2	89.8	1539	2	US-08-470-426B-17
6	431.2	89.8	1863	2	US-08-470-426B-14
7	428	89.2	2595	3	US-09-014-416-4
8	428	89.2	2599	3	US-09-014-416-6
9	426.2	88.8	633	3	US-08-612-973-7
10	426.2	88.8	633	3	US-08-927-597-7
11	423.4	88.2	501	2	US-08-483-695-30
12	423.4	88.2	501	2	US-07-965-285-30
13	423.4	88.2	501	2	US-08-487-231-30
14	423.4	88.2	501	3	US-09-201-912-30
15	422.8	88.1	9472	4	US-08-150-204E-36
16	422.2	88.0	642	3	US-08-612-973-3
17	422.2	88.0	642	3	US-08-927-597-3
18	421.2	87.8	795	3	US-08-612-973-5
19	421.2	87.8	795	3	US-08-927-597-5
20	421.2	87.8	2082	3	US-08-612-973-47
21	421.2	87.8	2082	3	US-08-927-597-47
22	421.2	87.8	2433	3	US-08-612-973-49
23	421.2	87.8	2433	3	US-08-927-597-49
24	419.6	87.4	1037	1	US-08-462-195-1
25	419.6	87.4	1037	2	US-08-636-883-1
26	419.6	87.4	1037	3	US-09-127-829-1
27	417.2	86.9	636	3	US-08-612-973-13

28 417.2 86.9 636 3 US-08-927-597-13 Sequence 13, Appl  
29 416.8 86.8 742 1 US-08-081-072-18 Sequence 18, Appl  
30 416.8 86.8 742 1 US-08-449-093A-18 Sequence 18, Appl  
31 416.8 86.8 932 1 US-08-081-072-15 Sequence 15, Appl  
32 416.8 86.8 932 1 US-08-449-093A-15 Sequence 15, Appl  
33 415.4 86.5 2116 3 US-08-191-160-21 Sequence 21, Appl  
34 413.8 86.2 501 2 US-08-483-695-28 Sequence 28, Appl  
35 413.8 86.2 501 2 US-07-965-285-28 Sequence 28, Appl  
36 413.8 86.2 501 3 US-08-487-231-28 Sequence 28, Appl  
37 413.8 86.2 501 3 US-09-201-912-28 Sequence 28, Appl  
38 413.6 86.2 11076 4 US-09-539-601-19 Sequence 19, Appl  
39 413.6 86.2 11076 4 US-09-539-601-19 Sequence 19, Appl  
40 413.6 86.2 11076 4 US-09-539-601-25 Sequence 25, Appl  
41 413.6 86.4 11076 4 US-09-539-601-31 Sequence 31, Appl  
42 405.2 84.4 1167 1 US-08-324-977-9 Sequence 9, Appl  
43 405.2 84.4 1167 2 US-08-384-616-9 Sequence 9, Appl  
44 405.2 84.4 1167 2 US-08-904-686A-9 Sequence 9, Appl  
45 405.2 84.4 1167 3 US-09-315-850-9 Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-08-612-973-11  
; Sequence 11, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARTYNOFF, GUY  
; APPLICANT: BUYSSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973  
; FILING DATE: 11-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BYRNE, THOMAS E.  
; REGISTRATION NUMBER: 32,205  
; REFERENCE/DOCKET NUMBER: 1487-10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 480 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..477  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..474



```

CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Review #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..480
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..477
US-08-612-973-9

```

Query Match	98.8%;	Score 474.2;	DB 3;	Length 483;
Best Local Similarity	99.4%;	Pred. No. 1.8e-125;		
Matches 476;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	ATGTCGGGTGCTCTTTCTCTATCTTCTCTCTGCGCCCTGCTGTGCTCTGTGACCATACCA	60	
Db	1	ATGCCCGGTTGCTCTTTCTCTATCTTCTCTGCGCCCTGCTGTGCTCTGTGACCATACCA	60	
Qy	61	GCTTTCGCTTATGAAGTGCACAGCTGTCGCGGGTGACCATGTACAGAAAGCTGCTCC	120	
Db	61	GCTTTCGCTTATGAAGTGCACAGCTGTCGCGGGTGACCATGTACAGAAAGCTGCTCC	120	
Qy	121	AACTCAAGCATAGTGTATGAGGACAGCGGACATGATCATGACACACCCCGGGTGCTGCC	180	
Db	121	AACTCAAGCATAGTGTATGAGGACAGCGGACATGATCATGACACACCCCGGGTGCTGCC	180	
Qy	181	TGGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGGTCACTCCACGCTCGCGGCC	240	
Db	181	TGGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGGTCACTCCACGCTCGCGGCC	240	
Qy	241	AGGAAGCAGAGCTCCCCCAACAGCAATACGACGCCACGTCGATTCGTTGGGGCT	300	
Db	241	AGGAAGCAGAGCTCCCCCAACAGCAATACGACGCCACGTCGATTCGTTGGGGCT	300	
Qy	301	GCTGCTTTCGTTCCGGCTATGTAGTGGGGGATCTCTGCGGATCTGTTTCCCTGTTTCC	360	
Db	301	GCTGCTTTCGTTCCGGCTATGTAGTGGGGGATCTCTGCGGATCTGTTTCCCTGTTTCC	360	
Qy	361	CAGCTGTTCACTTCTCAACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATC	420	
Db	361	CAGCTGTTCACTTCTCAACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATC	420	
Qy	421	TATCCCGGCCCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGAACTGGTAAATA	479	
Db	421	TATCCCGGCCCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGAACTGGTAAATA	479	

```

RESULT 4
US-08-927-597-9
; Sequence 9, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..480
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..477
US-08-927-597-9

Query Match 98.8%; Score 474.2; DB 3; Length 483;
Best Local Similarity 99.4%; Pred. No. 1.8e-125;
Matches 476; Conservative 0; Mismatches 3; Indels 0;

QY 1 ATGTCGGGTTCCTTTCTCTATCTTCTCTTGGCCCTGCTCTCTGTCTGACCA
DB 1 ATGCCCGGTTCCTTTCCTATCTTCTCTTGGCCCTGCTCTGTCTGACCA
QY 61 GCTTCGCTTATGAAGTCCGCAACGTTCGCGGGGTTCATGTACAGAACGACT
DB 61 GCTTCGCTTATGAAGTCCGCAACGTTCGCGGGGTTCACCATGTACAGAACGACT
QY 121 AACTCAACGATGCTATGAGCGAGCGGATGATGATCATGACACACCCCGGGTGGC
DB 121 AACTCAAGCATGTGTATGAGCGAGCGGATGATGATCATGACACACCCCGGGTGGC
QY 181 TCGCTTCGGGAGGGCAACTCCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCTCG

```

Db 181 TGGGTTGGGAGGCAACTCTCCCGTGTGGGTGGGCGCTCACTCCCAAGCTCGCGCC 240  
Qy 241 AGGAACGCCAGCGTCCCAACACGACAAATACGACGCGCCAGCTCGATTGCTGTTGGGCT 300  
Db 241 AGGAACGCCAGCGTCCCAACACGACAAATACGACGCGCCAGCTCGATTGCTGTTGGGCT 300  
Qy 301 GCTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGCGGATCTGTTTCTTTTCC 360  
Db 301 GCTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGCGGATCTGTTTCTTTTCC 360  
Qy 361 CAGCTGTTCACTTCTCACCTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC 420  
Db 361 CAGCTGTTCACTTCTCACCTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC 420  
Qy 421 TATCCCGGCCATGATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 479  
Db 421 TATCCCGGCCATGATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 479

RESULT 5  
US-08-470-426B-17  
; Sequence 17, Application US/08470426B  
; Patent No. 5856458  
; GENERAL INFORMATION:  
; APPLICANT: Okamoto, Hiroaki  
; APPLICANT: Nakamura, Tetsuo  
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR  
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B  
; TITLE OF INVENTION: HEPATITIS VIRUS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,  
; STREET: 1850 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,426B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-153402  
; FILING DATE: 12-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weillacher, Robert G.  
; REGISTRATION NUMBER: 20,531  
; REFERENCE/DOCKET NUMBER: 06/59-47083.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 659-2811  
; TELEFAX: (202) 659-1462  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1539 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
US-08-470-426B-17

Query Match 89.8%; Score 431.2; DB 2; Length 1539;  
Best Local Similarity 94.1%; Pred. No. 4.2e-113;  
Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2 TGGCGGTGCTCTTTCTCTATCTTCTTGGCTTTGCTTCTGTTGACCATCCAG 565

Db 506 TSCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTGCTTCTGTTGACCATCCAG 565  
Qy 62 CTTCCGCTTATGAAGTGGCAACGTTCTCGGCGTGTACCATGTGACGAACGACTCTCCA 121  
Db 566 CTTCCGCTTATGAAGTGGCAACGTTCTCGGCGTGTACCATGTGACGAACGACTCTCCA 625  
Qy 122 ACTCAAGCATAGTGTATGAGCGAGCGGACATGATGATGACACACCCCGGCTGCGGCT 181  
Db 626 ACTCAAGCATAGTGTATGAGCGAGCGGACATGATGATGATGATGATGATGATGAT 685  
Qy 182 GCGTTCGGGAGGCAACTCTCTCCGCTTCTGCTGGGTGGGCTCCTCACTCCAGCTCGGCGCA 241  
Db 686 GCGTTCGGGAGGCAACTCTCTCCGCTTCTGCTGGGTGGGCTCCTCACTCCAGCTCGGCGCA 745  
Qy 242 GGAACGCCAGCGTCCCAACACGACAAATACGACGCGCATGCTGCTTGGGCTG 301  
Db 746 GGAATGCCAGGCTCCCACTACGACAAATACGACGCGCATGCTGCTTGGGCTG 805  
Qy 302 CTGCTTTCTGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361  
Db 806 CTGCTTTCTGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865  
Qy 362 AGCTGTTCACTTCTCACCTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421  
Db 866 AGCTGTTCACTTCTCACCTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 925  
Qy 422 ATCCCGGCCATGATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGATGAT 477  
Db 926 ATCCCGGCCATGATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGATGAT 981

RESULT 6  
US-08-470-426B-14  
; Sequence 14, Application US/08470426B  
; Patent No. 5856458  
; GENERAL INFORMATION:  
; APPLICANT: Okamoto, Hiroaki  
; APPLICANT: Nakamura, Tetsuo  
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR  
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B  
; TITLE OF INVENTION: HEPATITIS VIRUS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,  
; STREET: 1850 M Street, N.W., Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,426B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-153402  
; FILING DATE: 12-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weillacher, Robert G.  
; REGISTRATION NUMBER: 20,531  
; REFERENCE/DOCKET NUMBER: 06/59-47083.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 659-2811  
; TELEFAX: (202) 659-1462  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1863 base pairs  
; TYPE: nucleic acid







```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..627
;
US-08-927-597-7

Query Match      88.8%; Score 426.2; DB 3; Length 633;
Best Local Similarity 93.1%; Pred. No. 8.4e-112;
Matches 446; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 2 TGTCCGGTGTCTTTCTCTATCTTCTCTCTTGGCCCTGCTGCTCTCTGACCATACCAG 61
Db 155 TGCCTGGTGTCTTTCTCTATCTTCTCTTGGCTTTACTGCTCTGTGACCATCCAG 214

Qy 62 CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGTGTACCATGTACGAAAGACTGCTCCA 121
Db 215 CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGTGTACCATGTACGAAAGACTGCTCCA 274

Qy 122 ACTCAAGCATAGTATGATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCCT 181
Db 275 ACTCAAGCATAGTATGATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCCT 334

Qy 182 GGTTCGGAGGAGCAACTCTCCCTGCTGCTGGGGTGTACCTCCAGCTGCGGGCCA 241
Db 335 GGTTCGGAGGAGCAACTCTCTCCCTGCTGCTGGGGTGTACCTCCAGCTGCGGGCCA 394

Qy 242 GGAAGCGCAGGCTGCCCAACAGCAATATACGACGCACTGATTTGCTGTTGGGGCTG 301
Db 395 GGAAGCGCAGGCTGCCCAACAGCAATATACGACGCACTGATTTGCTGTTGGGGCTG 454

Qy 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db 455 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 514

Qy 362 AGCTGTTCACTTCTACCTGCGCGGATGCAAGTACAGGACTGCAACTGCTCAATCT 421
Db 515 AGCTGTTCACTTCTACCTGCGCGGATGCAAGTACAGGACTGCAACTGCTCAATCT 574

Qy 422 ATCCCGGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 575 ATCCCGGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 633
```

## RESULT 11

```
US-08-483-695-30
; Sequence 30, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cDNA to genomic RNA
;
US-08-483-695-30
```

```
Query Match      88.2%; Score 423.4; DB 2; Length 501;
Best Local Similarity 94.4%; Pred. No. 4.9e-111;
Matches 439; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 2 TGTCCGGTGTCTTTCTCTATCTTCTCTCTTGGCCCTGCTGCTCTGACCATACCAG 61
Db 37 TGCCTGGTGTCTTTCTCTATCTTCTCTTGGCTTTGCTGCTCTGTTGACCATCCAG 96

Qy 62 CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGTGTACCATGTACGAAAGACTGCTCCA 121
Db 97 CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGTGTACCATGTACGAAAGACTGCTCCA 156

Qy 122 ACTCAAGCATAGTATGATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCCT 181
Db 157 ACTCAAGCATAGTATGATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCCT 216

Qy 182 GGTTCGGAGGAGCAACTCTCCCTGCTGCTGGGGTGTACCTCCAGCTGCGGGCCA 241
Db 217 GGTTCGGAGGAGCAACTCTCCCTGCTGCTGGGGTGTACCTCCAGCTGCGGGCCA 276

Qy 242 GGAAGCGCAGGCTGCCCAACAGCAATATACGACGCACTGATTTGCTGTTGGGGCTG 301
Db 277 GGAAGCGCAGGCTGCCCAACAGCAATATACGACGCACTGATTTGCTGTTGGGGCTG 336

Qy 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db 337 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 396

Qy 362 AGCTGTTCACTTCTACCTGCGCGGATGCAAGTACAGGACTGCAACTGCTCAATCT 421
Db 397 AGCTGTTCACTTCTACCTGCGCGGATGCAAGTACAGGACTGCAACTGCTCAATCT 456

Qy 422 ATCCCGGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 466
Db 457 ATCCCGGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
```

## RESULT 12

```
US-07-965-285-30
; Sequence 30, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
```



QY 122 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCAACCCCGGGTGGTGGCCCT 181  
DB 157 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCAATCTCCGGGTGGTGGCCCT 216  
QY 182 GCGTTGGGAGGCACTCTCCGTTGCTGGGTGGGCTCACTCCACGCTCGGGCCA 241  
DB 217 GCGTTGGGAGGCAACAGCTCCGTTGCTGGGTGGGCTCACTCCACGCTCGGGCCA 276  
QY 242 GGAAGCCAGCGTCCCAACACATACAGCCACGACGCTGATTTGCTCGTTGGGGCTG 301  
DB 277 GGAATGCCAGCGTCCCAACATACAGCCACGACGCTGATTTGCTCGTTGGGGCTG 336  
QY 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGGATCTGTTTCTCTGTTTCCC 361  
DB 337 CTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGGATCTGTTTCTCTGTTTCCC 396  
QY 362 AGCTGTTACCTTCTCACTCGCGGCAATCAACAGTACAGGACTGCAATGCTCAATCT 421  
DB 397 AGCTGTTACCTTCTCACTCGCGGCAATGAGACAGTGCAGGACTGCAATGCTCAATCT 456  
QY 422 ATCCCGGCAATGATCAGGTACCGCATGCGCTTGGGATATGATGA 466  
DB 457 ATCCCGGCAATGATCAGGTACCGCATGCGCTTGGGATATGATGA 501

## RESULT 14

US-09-201-912-30  
; Sequence 30, Application US/09201912  
; Patent No. 6210962  
; GENERAL INFORMATION:  
; APPLICANT: Brecht, Christian  
; APPLICANT: Krensdorf, Dina  
; APPLICANT: Porchon, Colette  
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a  
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic  
; TITLE OF INVENTION: Applications  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/201,912  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,285  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 05286-0001-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Other  
; DESCRIPTION: CDNA to genomic RNA

RESULT 15  
US-08-150-204E-96  
; Sequence 96, Application US/08150204E  
; Patent No. 6538126  
; GENERAL INFORMATION:  
; APPLICANT: CHO, Joong Myung  
; LEE, Young Beom  
; PARK, Young Woo  
; LIM, Kook Jin  
; CHOI, Deog Young  
; SO, Hong Seob  
; KIM, Chun Hyung  
; KIM, Sung Taek  
; YANG, Jae Young  
; TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES  
; NUMBER OF SEQUENCES: 128  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YANG, Jae Young  
; STREET: 386-1, Doryong-dong, Yuseong-gu  
; CITY: Daejeon  
; STATE: Daejeon  
; COUNTRY: Republic of Korea  
; ZIP: 305-340  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage  
; COMPUTER: IBM PC/pentium  
; OPERATING SYSTEM: Windows  
; SOFTWARE: Microsoft Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/150,204E  
; FILING DATE: 20-Apr-1994  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: KR 91-9510  
; FILING DATE: 10-JUN-1991

```

; APPLICATION NUMBER: KR 91-13601
; FILING DATE: 6-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahan Islam, Esq.
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER: 2695/FLK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 940-8564
; TELEFAX: (212) 940-8776
; INFORMATION FOR SEQ ID NO: 96
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: KHCV-LBC1, Fig. 2
; SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-204E-96

Query Match      88.1%; Score 422.8; DB 4; Length 9472;
Best Local Similarity 93.2%; Pred. No. 1.8e-110;
Matches 442; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      2 TGTGGGTGCTCTTCTCTATCTTCCTTGGCCCTGCTGCTCTGACCATACCAG 61
Db      848 TGCCTGGGTGCTCTTCTCTATCTTCCTTGGCTCTGCTCTTGTGTTGACCATCCAG 907
Qy      62 CTTCGGCTTATGAGTGGCAACGCTGTCGGGGGTGACCATCTCAGCAACGACTGCTCCA 121
Db      908 TTTTCGGCTTATGAGTGGCAACGCTGTCGGGGGTGACCATCTCAGCAACGACTGCTCCA 967
Qy     122 ACTCAAGCATAGTGTATGAGGAGAGCGGACATGATCATGACACACCCCGGGTGGTGCCT 181
Db     968 ACTCAAGCATAGTGTATGAGGAGAGCGGACATGATCATGACACACTCCCGGGTGGTGCCT 1027
Qy     182 GCGTTCGGGAGGCAACTCTCCCGTTCCTGGGTGGGCGCTCACTCCCAAGCTCGCGGCA 241
Db    1028 GCGTTCGGGAGGCAACTCTCCCGTTCCTGGGTGGGCACTTACTCCCAAGCTCGCGGCA 1087
Qy     242 GGAACGCCAGCGTCCCAACAGCAATACGACGCCACGTCGATTTGCTGGGGCTG 301
Db    1088 GGAATGCCAGCGTCCCACTAGCAATTCGACGCCCATGTGACTGCTCGTGGGGTAG 1147
Qy     302 CTGCTTTCTGTTCCGCTATGTATGATGCGGGGATCTTCGGGATCTGTTTCTTGTTCCTCC 361
Db    1148 CTGCTTTCTGTTCCGCTATGTATGATGCGGGGACCTCTGCGGATCTGTTTCTTGTTCCTCC 1207
Qy     362 AGCTGTTACCTTCTACCTCGCGGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db    1208 AGCTGTTACCTTCTACCTCGCGGGCATGAGACGGTACAGGACTGCAACTGCTCAATCT 1267
Qy     422 ATCCCGGCCATGATACAGTACCGCATGGCTTGGGATATGATGATGATGATGATGAT 475
Db    1268 ATCCCGGCCGATATAGGTACCGCATGGCTTGGGATATGATGATGATGATGATGAT 1321
```

Search completed: June 16, 2004, 13:30:24  
Job time : 35.6 secs